UNIVERSITY OF WISCONSIN-LA CROSSE

Graduate Studies

THE EFFECTS OF GENETIC BOTTLENECKS ON MUTATION FIXATION AND REPLICATIVE CAPACITY OF THE INFLUENZA A VIRUS

A Manuscript Style Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Clinical Microbiology

Michael Mamerow

College of Science and Health

Clinical Microbiology

December, 2018

THE EFFECTS OF GENETIC BOTTLENECKS ON MUTATION FIXATION AND REPLICATIVE CAPACITY OF THE INFLUENZA A VIRUS

By Michael Mamerow

We recommend acceptance of this thesis in partial fulfillment of the candidate's requirements for the degree of Master of Science in Clinical Microbiology

The candidate has completed the oral defense of the thesis.

Peter Wilker, Ph.D., Thesis Committee Chairperson — 11-8-18 Date

Michael Hoffman, Ph.D., Thesis Committee Member — 11-8-18 Date

Marc Rott, Ph.D., Thesis Committee Member — 11/8/18 Date

Kelly Gorres, Ph.D., Thesis Committee Member — 11/8/18 Date

Thesis accepted

Meredith Thomsen, Ph.D., Graduate Studies Director — 11-27-2018 Date

ABSTRACT

Mamerow, M.A. The effects of genetic bottlenecks on mutation fixation and replicative capacity of the influenza A virus. MS in Clinical Microbiology, December 2018, 64pp (P, Wilker)

The influenza A virus is a common cause of respiratory illness in humans. Seasonal epidemics of influenza in the United States can result in up to 40 million cases, with annual hospitalizations and deaths reaching as high as 400,000 and 70,000, respectively. The influenza A virus continually causes annual epidemics, despite yearly vaccines, as a result of its high mutation rate, leading to genetically diverse viral populations within hosts. Respiratory droplet-mediated transmission of the virus between individuals is accompanied by a bottleneck event that decreases the diversity of the viral population passed to new hosts. Using controlled artificial bottlenecks of different sizes during *in vitro* serial passaging, the impact of these bottleneck events on patterns of mutation fixation and replicative capacity of the influenza A virus was determined. Growth curves and genome sequencing were used after passaging to characterize differences in replicative capacity and identify genomic changes within the population. Serial passaging at a bottleneck size of one virus generated virus populations with lower replicative capacity as compared to the original parental virus. This effect was associated with fixation of numerous mutations spread throughout the genome. An increase in replicative capacity was evident after repeated bottlenecks of 1000 viruses, demonstrating that sufficiently loose bottlenecks do not compromise viral replication kinetics and even allow for improvement. This increase in replicative capacity was associated with several mutations, clustered primarily in the hemagglutinin genome segment.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

INTRODUCTION

Influenza Background

The family *Orthomyxoviridae* contains three genera of influenza viruses – influenza A, B, and C. These are single-stranded, negative-sense RNA viruses with segmented genomes containing seven or eight genome segments. Influenza A, in particular, contains eight genomic segments spanning roughly 14,000 nucleotides (1, 2). Characterization of influenza viruses into type A, B, or C is most commonly done based on the serological response to internal proteins, specifically the matrix (M) and nucleoprotein (NP) proteins (3). Influenza A can be further separated into subtypes based on differences in the surface proteins hemagglutinin (HA) and neuraminidase (NA). Currently, 18 HA and 11 NA subtypes have been discovered, though only subtypes H1N1 and H3N2 currently circulate amongst human populations (4). Both influenza B and C viruses are primarily isolated from humans, though type B has been found in seals, and type C in pigs and dogs (2). Wild aquatic birds are the natural reservoir for most subtypes of influenza A, the exceptions being HA subtypes 17/18 and NA subtypes 10/11 which are only found in bats (4). Influenza A virus also causes disease in a wide variety of warm-blooded animals. Though only type A viruses have been seen to cause global pandemics, both influenza A and B are capable of causing large-scale epidemics, with such outbreaks often occurring in a seasonal fashion, particularly in the winter months (2, 5, 6). Type A viruses also cause the most severe disease, while type C viruses are the

least severe, usually being limited to a mild respiratory illness similar to the common cold (3).

Infections with influenza A virus (IAV) result in a respiratory illness that can affect both the upper and lower respiratory tracts. As a respiratory illness, influenza is primarily transmitted from person to person through respiratory droplets, though transmission through fomites is also possible. Persons infected with the influenza virus are often contagious 24 hours before symptoms appear and remain contagious for five days (7). Influenza can manifest with a variety of non-specific symptoms including fever, chills, cough, sore throat, and muscle or body aches. In most healthy individuals, the disease lasts for one to two weeks with no complications. However, in young children, the elderly, and those with underlying medical conditions, IAV may produce a much more serious illness that can result in moderate to severe complications, including pneumonia (both primary viral and secondary bacterial), hemorrhagic bronchitis, and sepsis (2, 8).

Significance of Influenza

The impact of seasonal influenza is exceptionally hard to determine, due to variability in the severity and number of cases between seasonal outbreaks. However, between medical costs, lost earnings or loss of life, it has been estimated that the total economic burden of seasonal outbreaks of influenza in the United States is $87.1 billion annually (9). In the last five years, estimates for the total number of cases have ranged from 7 million to 40 million in the US (10, 11). The better measures for the burden of influenza are the number of influenza-associated hospitalizations and deaths. During seasonal outbreaks in the US from 1979-2001, estimates for the annual number of

influenza-associated hospitalizations ranged from 55,000 to 431,000, with a mean of 226,000 (10, 11). During that same period, the estimated annual number of influenza-related deaths ranged from 3,300 to 70,000 (10, 11). Though the influenza virus can cause disease in people of any age, the majority of hospitalizations and deaths due to influenza are in those aged 65+ (11, 12). Emergence of pandemic influenza can greatly increase the overall burden of influenza, both through an increase in the number of cases and severity of the disease. The Spanish influenza pandemic of 1918 infected an estimated 500 million people, one third of the entire global population at the time, leading to roughly 50 million deaths. The mortality rate of the 1918 pandemic was 25 times greater than the average mortality rate of season influenza (13). The 2009 H1N1 'swine flu' pandemic, while not resulting in excessive mortality relative to seasonal epidemics, caused disease in younger people much more effectively than seasonal influenza, possibly as a result of lower immunity to H1N1 viruses in this age group. Hospitalization and death rates in children and working adults were four to seven and eight to twelve times greater, respectively, due to 'swine flu' compared to rates normally caused by seasonal influenza viruses (14). The continued burden of seasonal epidemics and the sudden appearance of pandemic strains are due in part to the potent ability of influenza to change over time.

Influenza Mechanisms of Variation

Antigenic drift is the process through which viruses accumulate genomic changes that lead to amino acid substitutions within their antigenic sites. In all viruses, influenza included, antigenic drift occurs as a product of error-prone genome replication and natural selection. Viruses harboring beneficial mutations can increase in frequency within

a population through selective pressure, primarily the human immune response. Host antibodies against the influenza viral particle promote evolution of the virus – specifically within the genome segments coding for HA or NA surface proteins (15). Drift occurring in these genome segments can lead to changes in the antibody binding sites of the proteins, reducing the effectiveness of pre-existing antibodies (10, 16). New strains of influenza can develop due to these antigenic changes, which is one of the reasons vaccines must be periodically updated and the strains included in the vaccine carefully selected. Hemagglutinin in particular has evolved in a very distinct stepwise fashion over the last several decades as a direct result of selective pressure on the influenza virus. Since the introduction of the H3N2 subtype into the human population in the 1968 Hong Kong influenza pandemic, evolution of the H3 gene has occurred in a single lineage. Antigenically different H3 proteins have emerged roughly every 1-2 years, based on data used for vaccine recommendations (17). Evolution of influenza – through antigenic drift and driven by selective pressure of the host immune system – is a major factor contributing to the seasonal nature of outbreaks of the virus but does not explain the sudden appearance of highly virulent pandemic strains.

Pandemic strains of influenza primarily arise as a result of antigenic shift, a type of genetic reassortment that occurs in influenza A, but not B or C. Genetic reassortment is defined as the exchange of genetic material between two distinct viruses infecting the same cell. In influenza A, this process is facilitated by its segmented genome, which allows for the efficient reassortment of entire genome segments between different viruses. Antigenic shift is a specific occurrence of genetic reassortment between subtypes of influenza, primarily involving the genome segment for the HA protein, as the HA

protein is a primary target of protective antibody responses, though other segments can be involved. This leads to the introduction of novel subtypes of HA and NA into human populations, against which there is little existing immunity (2). The emergence of pandemic strains often involves reassortment of subtypes from different animals. However, both avian subtypes and human subtypes exhibit a high degree of host-specificity. Pigs, on the other hand, can be efficiently infected by both avian and human influenza viruses, making them effective intermediates in which antigenic shift can happen. The lack of host specificity seen in pigs is due to the nature of the sialic acid residues, which act as the receptor for influenza A viruses, found on their cells.

The influenza virus binds to sialic acid residues present on host epithelial cells in order to enter into the cell. These residues are linked to galactose residues at the end of carbohydrates on cell surfaces. The sialic acid residue can be linked to galactose in two ways; the carbon atom at position number 2 of the sialic acid hexose can be joined to the carbon at position 3 of the hexose of galactose (α-2,3 linkage) or to the carbon at position 6 of the galactose (α-2,6 linkage) (18, 19). Avian HA subtypes preferentially bind to sialic acid connected to galactose via an α-2,3 linkage, while human subtypes preferentially bind those connected via an α-2,6 linkage. Cells within the pig respiratory tract contain both types of sialic acid linkages, allowing them to be infected by both avian and human viruses. Furthermore, the presence of both of these linkages may allow for the adaptation of avian viruses to the α-2,6 linkage, which is dominant in the human respiratory tract, potentially leading to the emergence of avian-derived viruses that can infect humans (17). The adaptation of novel and avian viruses in this manner may lead to

more efficient transmission between humans and aid the virus in reaching its true pandemic potential (2).

Influenza is constantly evolving, even beyond antigenic shift and drift, primarily due to being an RNA virus. The RNA-dependent RNA polymerase of influenza transcribes its genome into positive sense mRNA (for translation into protein) or cRNA (for replication of the genome). The polymerases of RNA viruses contain no proofreading ability, resulting in error rates ranging from 10^{-4} to 10^{-5} substitutions per site, though IAV has been measured to have rates as high as $8x10^{-3}$ (2, 17). With a genome merely 14 kb long, this introduces, on average, one mutation into the genome as a whole each time it is replicated into cRNA. High rates of mutation, coupled with high replication rates and large population sizes generated during infection, leads to very large and very diverse viral populations (20). Over the course of an infection in a single individual, it is estimated that every possible point mutation occurs within these large viral populations. This population is known as a quasispecies: a collection of closely related but ultimately distinct viruses organized around a master sequence, where the master sequence is defined as being the most fit variant selected for amongst the quasispecies (20–22, 24). Genetic diversity, or the number and frequency of variants within a quasispecies, is heavily reliant on the tolerance to mutations produced during replication (24). Viral diversity within a population is constantly in flux, being impacted by continuous competition and selection, with the fitness of viral progeny determining their frequency within a population. In general, positive selection allows for the dominance of higher fitness clones while eliminating those with lower fitness. In contrast to the constant nature of competition, genetic bottlenecks are distinct events that can both

specifically and stochastically alter viral diversity. These genetic bottlenecks, as must occur during transmission, can dramatically reduce the population size and may allow only a small subset of variants, with varying degrees of fitness, to survive and replicate (25).

Genetic Bottlenecks and the Influenza Virus

Influenza can be transmitted person-to-person in a number of ways. The consensus is that the most common means through which influenza spreads from person to person is through large respiratory droplets, defined as respiratory particles larger than 5 μm in diameter. These droplets are generated through coughing, sneezing or talking, do not remain suspended in the air, and require close contact to successfully spread disease (26, 27). The virus can also become aerosolized through coughing or sneezing and transmission can occur as microscopic droplets, often smaller than 5 μm, remain suspended in the air for long periods of time (26). Additionally, influenza may be spread directly, through physical contact with respiratory secretions from an infected host, and indirectly, through contact with contaminated surfaces or objects in the environment.

Both aerosolized and large respiratory droplets can contain anywhere from 1 to >1000 viral particles, though droplets typically contain under 100 particles (28–30). Within infected individuals, viral populations can be extremely large – populations in the millions are not uncommon, and some viral populations can be as large as 10^{12} particles per human, and influenza is no exception (24, 31). Thus, transmission events in which merely 100 virus particles are passed on from a population of millions or billions represent a very significant genetic bottleneck. These events are defined by a sharp decrease in a population, leading to a significant reduction in the population diversity. In

viral populations, genetic bottlenecks, especially repeated bottlenecks, are often accompanied by a loss of fitness (31, 32). This should not come as a surprise – if most mutations are deleterious, then most progeny are likely less fit than the parent from which they were generated. The chances of selecting high fitness viruses in the event of a genetic bottleneck are very small, leading to less fit viruses being spread to new hosts.

Previous work on the influenza virus supports the idea that transmission events present significant genetic bottlenecks, reducing the diversity of the viral population that is passed on to new hosts. In one such study, researchers produced over 100 molecularly barcoded influenza viruses by inserting unique, detectable 22 nucleotide barcode identifiers into the NS genome segment. Through sequencing of the barcode identifiers before and after transmission, dynamics of transmission-associated bottlenecks could be determined. These barcoded viruses were ultimately used to determine that airborne transmission of the virus in an animal model produced very significant genetic bottlenecks, with some new infections being initiated by as few as two barcoded viruses, from an initial starting population of over 100. Additionally, the probability of any given variant in a population to be passed on through the bottleneck was found to be solely dependent on its proportion within the original population (33). Other work tracking avian HA haplotypes passed to new hosts in a ferret model also found that transmission is accompanied by a bottleneck. A diverse population of HA haplotypes in experimentally infected index ferrets was reduced to a much smaller number of haplotypes in newly infected ferrets, suggesting this transmission was mediated by a very small subset of the original population (34). Though a radically different virus than influenza A, HIV has also shown evidence for bottleneck events during transmission. New infections were

observed to be initiated by a single viral particle, representing a very drastic reduction in the population being passed on to new hosts (35). This previous work supports the idea that transmission events, whether in influenza or other viruses, lead to a bottleneck of the viral population and a reduction of the diversity of viruses transferred to new hosts.

Muller's Ratchet Theory

The steady loss of fitness after repeated genetic bottlenecks can be explained by Muller's ratchet. More specifically, Muller's ratchet describes a finite, asexual population that accumulates irreversible deleterious mutations, leading to a loss in fitness (36). Due to high mutation rates, eventually every member within a population will acquire a mutation. As an asexual population, no genetic recombination can occur that may be able to reverse these mutations and the occurrence of revertant and beneficial mutations are so low as to not be significant. Once every member of the population has acquired a single mutation, the ratchet "clicks" and the overall fitness of the population is irreversibly lowered (37). As this continues, the least-loaded class within the population – those individuals with the lowest number of mutations – acquires more and more mutations, potentially leading to extinction of the population (38). Muller's ratchet can occur naturally in a sufficiently small population, where a fitness equilibrium cannot be reached before all members of a population accumulate mutations, but is certainly accentuated through repeated bottleneck events (37, 39).

Though it began as a hypothesis, Muller's ratchet has been observed in populations of RNA viruses exposed to repeated genetic bottlenecks. Repeated plaque-to-plaque transfers of bacteriophage Φ6, vesicular stomatitis virus (VSV), and foot-and-mouth disease virus (FMDV) led to significant losses in fitness (39–41). Plaque-to-plaque

transfers represent a bottleneck size of a single particle, intensifying genetic drift that occurs naturally in RNA viruses (41). Interestingly, these RNA viruses undergo varying levels of recombination and reassortment, which could potentially allow the viruses to elude fitness losses from mutation. VSV and FMDV are single-stranded viruses, while Bacteriophage Φ6 contains a segmented genome, like influenza, allowing for reassortment between segments. Reassortment can potentially eliminate mutations, if two co-infecting viruses have mutations on separate segments. However, in experiments with the bacteriophage, mutations were not limited to a single segment, rendering reassortment ineffective at circumventing deleterious mutations (41). FMDV, a non-segmented RNA virus, cannot undergo reassortment, but exhibits high rates of recombination during replication (39). Despite this, plaque-to-plaque transfers led to a significant loss in fitness in the majority of passaged clone populations (39). VSV is also a non-segmented RNA virus, but unlike FMDV does not undergo measurable recombination during replication. Unsurprisingly, fitness losses were also observed in most clone populations after repeated passages of VSV (40). Though direct comparisons between experiments with FMDV and VSV are tenuous at best, it is interesting that plaque-to-plaque transfers had a similar effect on virus populations in both the presence and absence of recombination. This would suggest that recombination, and even reassortment, may not be sufficient to counteract the effects of Muller's ratchet.

Though Muller's ratchet remains irreversible in small and bottlenecked populations, further work on VSV showed that large-scale passages of greater than 100,000 viruses were enough to not only return fitness, but improve it beyond original levels (40, 42). Bottleneck events prevent natural selection from removing low fitness

mutants whereas large-scale passages bring selection back into the fold, allowing higher fitness clones to out-compete other mutants. Additionally, large-scale passages increase the chances that clones with advantageous mutations are among the new population that is passed on. It is important to note, however, that the potential fitness gains from large population passages are highly dependent on the fitness of the population being passaged. In a very high fitness population, the likelihood of mutations occurring that further increase fitness is very low. Thus, even high population passages may fail to provide enough "better" clones that could lead to an increase in fitness. On the other hand, if the fitness of the population is very low, then mutations are much more likely to be advantageous. Large-scale passages of such a population would contain many viruses with higher fitness than the population mean, leading to very drastic and rapid fitness recovery (43).

Studies looking at the effect of genetic bottlenecks have been done on both segmented RNA viruses (bacteriophage Φ6) and single-stranded, negative sense RNA viruses (VSV), yet it remains unclear how much overlap these viruses have with influenza, a segmented, negative sense, single stranded, RNA virus that is repeatedly subjected to poorly defined bottleneck events during respiratory droplet-mediated transmission. A segmented genome appears to be insufficient in counteracting Muller's ratchet in drastic bottleneck events. However, influenza A has eight segments compared to three segments with bacteriophage Φ6, which theoretically means there needs to be a much greater accumulation of mutations before reassortment cannot recoup fitness losses.

Research Objectives

Here we sought to define the impact of serial genetic bottleneck events of varied sizes on the replicative capacity and mutational pattern of influenza viruses using an *in vitro* model system. We used a system of controlled artificial bottlenecks during *in vitro* serial passage of an influenza virus to model the effects of various types of genetic bottlenecks that have been shown to occur during aerosol-mediated transmission between mammals. Seasonal epidemics of influenza must involve significant amounts of repeated bottleneck events as the virus transmits between individuals. However, extinction of the virus as a result of fitness losses after transmission events does not occur, despite this being an expected outcome of repeated genetic bottlenecks. Loose bottlenecks that may circumvent Muller's ratchet could allow viruses to avoid these extinction events. This work hopes to shed some light on the exact impact of these bottleneck events and how similar events during seasonal epidemics actually affect the influenza virus. Additionally, this work could be useful in extrapolating the role of bottlenecks in the adaptation of avian influenza viruses to humans. Work on avian influenza viruses have shown that as few as four or five mutations are necessary to allow airborne transmission of these viruses in mammalian models (44, 45). In addition to the accumulation of these mutations, any avian viruses must also survive a bottleneck as they are transmitted from avian to mammalian hosts. Working to characterize these bottleneck events could provide insight into the type of obstacle these events present to avian influenza viruses. In order to accomplish our goal of determining the impact of serial bottlenecks on the influenza A virus, we defined two specific aims, which are outlined below:

Aim 1. Generate lines of influenza A/Victoria/361/2011 (H3N2) that have been subjected to repeated genetic bottlenecks of defined size during serial passage *in vitro*.

Aim 2. Compare the replicative capacity and genomic sequence of serially 'bottlenecked' virus lines to the parental virus.

MATERIALS AND METHODS

Experimental Overview

To determine the effect of genetic bottleneck size on replicative capacity, influenza A/Victoria/361/2011 (H3N2) was subjected to artificial genetic bottleneck sizes of 10^0 and 10^3 viruses during serial passage in cell culture (Fig. 1). This specific parental strain of influenza was chosen to minimize risk of human infection, as it was the primary target of the 2011 influenza vaccine. Thirty plaque-to-plaque transfers were done to produce viral lines subjected to a bottleneck size of 1 virus. For a bottleneck size 10^3 viruses, each round of sequential growth in culture was initiated using the number of viruses corresponding to the bottleneck size. Three independent lineages were generated for each bottleneck size. Harvested virus from every passage number, or plaque, in the case of the plaque-to-plaque transfers, was saved and frozen and plaque assays were done to determine the concentration of harvested viruses after each round of growth so that the next round could be initiated with the proper number of infectious virions.

Determination of Viral Concentration by Plaque Assay

Twelve-well plates were seeded one day prior to the assay with 1 ml/well of Madin-Darby Canine Kidney (MDCK) cells, a cell line routinely used for influenza work, at a concentration of 2.25×10^5 cells/ml and incubated at 37°C with 5% CO_2. Each well containing cells was washed twice with phosphate buffered saline (PBS) after which the

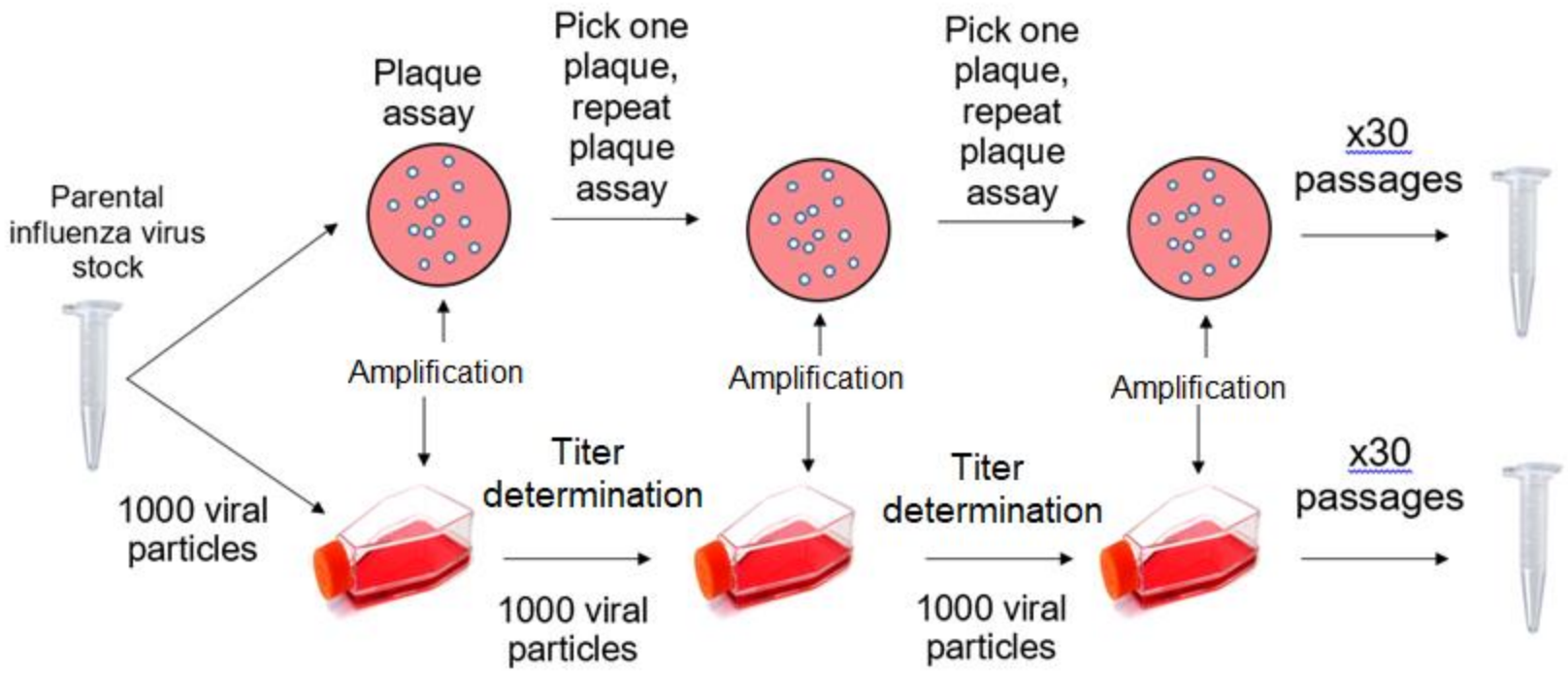

FIGURE 1 Experimental design of serial bottlenecks of 1 and 1000 viral particles. Parental virus A/Victoria/361/2011 (H3N2) was used to perform a routine plaque assay using MDCK cells, after which virus was harvested from a plaque and used for subsequent plaque assays, out to 30 passages, for the plaque-to-plaque transfers. For a bottleneck size of 10^3 viruses, parental virus was used to infect MDCK cells in a T25 flask for 3 days, harvested, frozen for one day, and then used to infect a new flask of MDCK. This process was repeated 30 times for the serial passaging process.

wells were inoculated with 200 µl of viral dilutions, with any control wells receiving 200 µl of Minimal Essential Media (MEM) supplemented with bovine serum albumin (BSA), sodium bicarbonate, MEM amino acids, antibiotics, an antimycotic, ThermoFischer MEM 100x vitamin solution (Cat # 11120052), and HEPES buffer, herein referred to as MEM-BSA. For the virus being quantified, a ten-fold dilution series was prepared in MEM-BSA out to a final dilution determined reasonable for the virus being used. After inoculation, plates were incubated for 45 minutes at 37°C with 5% CO_2. After incubation, viral inoculant was aspirated off and each well was washed again with PBS. Cells then received a 1 ml overlay of MEM-BSA containing 1% molten SeaPlaque agar, and TPCK-treated trypsin at a final concentration of 1 µg/ml. Once the overlay had solidified, the plates were incubated upside-down at 37°C and 5% CO_2 for three days. After the three-day incubation period, plates were removed and fixed for a minimum of four hours with 10% buffered formalin and stained with crystal violet dye. Plaques were counted to determine the concentration of virus harvested from the infected flasks using the following equation:

$$\text{Concentration (PFU/ml)} = \frac{\text{Average of 2 plaque counts}}{(\text{inoculant volume})(\text{dilution factor})}$$

Plaque-to-Plaque Transfer

Viral stocks were diluted 10-fold in MEM-BSA. Six-well plates were seeded one day prior to infection with 2 ml/well of MDCK cells at a concentration of roughly 2.25×10^5 cells/ml and incubated at 37°C with 5% CO_2. Wells were washed twice with PBS and then inoculated with 400 µl of the viral dilutions. Infected plates were incubated at 37°C with 5% CO_2 for 45 minutes with occasional rocking. After incubation,

remaining viral inoculant was aspirated off and all wells washed with PBS. A solid overlay was prepared in an identical manner as the overlay for plaque assays and 2 ml was added to each virally inoculated well and left to solidify, at which point the plates were incubated at 37°C and 5% CO_2 for three days. Once the three-day incubation period had passed, plaques were visualized by holding plates up to a light and circled on the underside of the plate. Circled plaques were picked using a 20 µl pipette and dispensed into 500 µl of MEM-BSA. Picked plaques were diluted out in 500 µl PBS and used for a dilution series in order to start the next plaque assay, repeating the process.

1000 PFU Transfer

Three T25 flasks, one flask for each viral line, were seeded one day prior to infection with 3.5 ml of cells at a concentration of $2.25x10^5$ cells/ml. Using viral concentrations obtained from plaque assays, virus from each of the three viral lines was diluted out to produce a solution of virus at a concentration of 1000 plaque-forming units (PFU)/ml. The following day, cells were washed twice with PBS and inoculated with 1 ml of the freshly prepared, 1000 PFU/ml viral solution. Infected flasks were incubated at 37°C, 5% CO_2 for 45 minutes with occasional rotation, after which all remaining liquid was aspirated off and replaced with 3.5 ml of MEM-BSA containing TPCK-treated trypsin at a final concentration of 1 µg/ml. After infection with the virus, flasks were incubated at 37°C, 5% CO_2 for three days at which point 90% of cells were dead and lifted from the bottom of the flasks. Media from infected flasks was pipetted into a 15 ml tube and centrifuged at 2,297 x G for five minutes to pellet cell debris. Supernatant was removed, placed into new 15 ml tubes, and vortexed for ~10 seconds. Finally, 300 µl

aliquots of the supernatant were pipetted into sterile 1.5 ml microfuge tubes, labeled with the passage number, bottleneck size, and date and stored at -70°C.

Multi-Cycle Growth Curve

Four T25 flasks, one flask for each viral line and one for the parental virus, were seeded one day prior to infection with 3.5 ml of cells at a concentration between 2.0 and $2.5x10^5$ cells/ml. The concentration, determined through manual counting using a hemocytometer, was recorded and the total number of cells in the flask 24 hours (when confluency was 90-100%) after seeding was estimated using the following equation:

$$Total\ cells = Cell\ concentration\ at\ seeding\ x\ volume\ added\ to\ flask\ x\ 2$$

Using viral concentrations obtained from plaque assays, virus from each of the three viral lines was used to infect the flask at a multiplicity of infection (MOI) of 0.001. For infection, cells were washed twice with PBS and inoculated with 1 ml of the freshly prepared viral stocks. Infected flasks were incubated at 37°C, 5% CO_2 for 45 minutes with occasional rotation, after which all remaining liquid was aspirated off and replaced with 3.5 ml of MEM-BSA containing TPCK-treated trypsin at a final concentration of 1 μg/ml. After infection with the virus, flasks were incubated at 37°C, 5% CO_2 for three days. Samples were taken every 12 hours. For the 12-hour sample, two aliquots of 110 μl from each flask were taken, placed into a 1.5 ml tube, and frozen. All other time points consisted of two 30 μl aliquots removed and frozen in the same manner. Plaque assays run in duplicate with parallel dilution series for each virus at each time point were performed in order to determine the concentration of the virus at the respective time point.

Viral Amplification

Three T75 flasks, one flask for each viral line, were seeded one day prior to infection with $2.25x10^6$ cells. The total number of cells in the flask 24 hours after seeding was estimated in the same manner as for the multi-cycle growth curve. The following day, virus from each of the three viral lines was used to infect cells at an MOI of 0.001 for the 1000 PFU viruses or an MOI of 0.01 for the plaque-to-plaque viruses. Different MOIs were used to account for differences in replicative capacity. Infected flasks were incubated at 37°C, 5% CO_2 for 45 minutes with occasional rotation, after which all remaining liquid was aspirated off and replaced with 10 ml of MEM-BSA containing TPCK-treated trypsin at a final concentration of 1 μg/ml. Flasks were then incubated at 37°C, 5% CO_2 for 24 or 36 hours, for the 1000 PFU viruses and plaque-to-plaque viruses, respectively. Following the incubation period, the virus was harvested using the same protocol as described in the "Virus Recovery" sub-section.

Single-Cycle Growth Curve

Four T25 flasks, one flask for each viral line and one for the parental virus, were seeded one day prior to infection with 3.5 ml of cells at a concentration between 2.0 and $2.5x10^5$ cells/ml. The total number of cells in the flask 24 hours after seeding was estimated using the same protocol as described in the "Multi-Cycle Growth Curve" subsection. Stocks of amplified virus from each of the three viral lines were prepared to produce an MOI of 5, based on the estimated number of cells. Preparation and infection of the flasks again followed the previously described protocol. After infection with the virus, flasks were incubated at 37°C, 5% CO_2 for 12 hours. Two 30 μl aliquots were taken every three hours, placed into a 1.5 ml tube, and frozen. Plaque assays run in

duplicate with parallel dilution series for each virus at each time point were performed in order to determine the concentration of the virus at the respective time point.

RNA Purification and Reverse Transcription

Viral RNA was purified from the viral stocks using the Ambion PureLink™ RNA Mini Kit (Cat # 12183018A) after which the RNA was reverse transcribed into cDNA using the Invitrogen SuperScript III First-Strand Synthesis System for RT-PCR. For reverse transcription, the primer AUni12F (5'-AGCRAAAGCAGG-3'), targeting 12 bases in the upstream untranslated region that are highly conserved amongst all influenza A viruses, was used. Both RNA purification and reverse transcription (RT) followed the procedures described by the respective kits, including thermocycler parameters.

Sequencing

Viral genome segments were amplified individually using the BioRad iProof™ High Fidelity PCR kit (Cat # 1725330). Samples were prepared following the kit-provided recommendations for 50 µl reactions, with the exception of $MgCl_2$, which was added to reach a final concentration of 3.4 mM. Additionally, 2 µl of viral cDNA was added as template to each reaction. Primers were added individually to each reaction, at a volume of 2.5 µl, to reach a final concentration of 5 µM. Primer sets and the region of each genome segment amplified by said primers for each passaged viral line and the parental virus can be seen in Table 1. Those primers labeled as M13F or M13R include a tag to facilitate downstream sequencing using the M13 forward (M13F, underlined in Table 2), or M13 reverse (M13R, bolded in Table 2) primers as described in Table 2. Primers not containing these M13 tags were sequenced using the same primers used in amplification. PCR amplification followed one of two thermocycler protocols. The first

TABLE 1 Primer pairs used for segmented genome amplification.

Primer	Bases Amplified	Parental	Plaque-to-plaque			1000 PFU		
			A30	B30	C30	A30	B30	C30
A-PB2-I-M13F A-PB2-I-M13R	1-864	✓				✓	✓	✓
A-PB2-II-M13F A-PB2-II-M13R	745-1694	✓	✓		✓	✓	✓	✓
A-PB2-III-M13F A-PB2-III-M13R	1469-2340	✓	✓	✓	✓	✓	✓	✓
P30-PB2-I-F P30-PB2-I-R	1-787							
P30-PB2-II-F P30-PB2-II-R	625-1317			✓	✓			
P30-PB2-III-F P30-PB2-III-R	1101-1785			✓	✓			
P30-PB2-I-F A-PB2-I-M13R	1-864		✓	✓	✓			
A-PB1-I-M13F A-PB1-I-M13R	1-820	✓	✓			✓	✓	✓
A-PB1-II-M13F A-PB1-II-M13R	693-1634	✓	✓			✓	✓	✓
A-PB1-III-M13F A-PB1-III-M13R	1520-2334	✓	✓	✓	✓	✓	✓	✓
P30-PB1-II-F P30-PB1-II-R	596-1350			✓	✓			
P30-PB1-III-F P30-PB1-III-R	1225-1817			✓	✓			
A-PB1-II-M13F P30-PB1-II-R	693-1350			✓	✓			
A-PA-I-M13F A-PA-I-M13R	1-980	✓	✓		✓	✓	✓	✓
A-PA-II-M13F A-PA-II-M13R	861-1619	✓	✓	✓		✓	✓	✓
A-PA-III-M13F A-PA-III-M13R	1435-2231	✓	✓		✓	✓	✓	✓
P30-PA-I-F P30-PA-I-R	6-710			✓				
P30-PA-II-F P30-PA-II-R	590-1343			✓	✓			

TABLE 1 (Continued)

Primer	Bases Amplified	Parental	Plaque-to-plaque			1000 PFU		
			A30	B30	C30	A30	B30	C30
P30-PA-III-F P30-PA-III-R	1137-1829			✓	✓			
P30-PA-IV-F P30-PA-IV-R	1689-2221			✓				
A-HA-I-M13F A-HA-I-M13R	5-1199	✓	✓	✓	✓	✓	✓	✓
A-HA-II-M13F A-HA-II-M13R	843-1763	✓	✓			✓	✓	✓
P30-HA-II-F P30-HA-II-R	1351-1724			✓	✓			
P30-HA-II-F A-HA-II-M13R	1351-1763			✓	✓			
A-NA-I-M13F A-NA-I-M13R	1-845	✓		✓	✓			
A-NA-II-M13F A-NA-II-M13R	723-1467	✓	✓			✓	✓	✓
NA-I M13F v2 NA-I M13R v2	8-828					✓	✓	✓
P30-NA-I-F P30-NA-I-R	1-608		✓		✓			
P30-NA-II-F P30-NA-II-R	350-1034		✓	✓	✓			
P30-NA-III-F P30-NA-III-R	819-1411			✓	✓			
P30-NA-III-F A-NA-II-M13R	819-1467			✓	✓			
A-NP-I-M13F A-NP-I-M13R	9-1026	✓				✓	✓	✓
A-NP-II-M13F A-NP-II-M13R	841-1568	✓	✓	✓	✓	✓	✓	✓
P30-NP-I-F P30-NP-I-R	2-613		✓	✓	✓			
P30-NP-II-F P30-NP-II-R	407-1026		✓	✓	✓			
A-NS-I-M13F A-NS-I-M13R	1-891	✓						

TABLE 1 (Continued)

Primer	Bases Amplified	Parental	Plaque-to-plaque			1000 PFU		
			A30	B30	C30	A30	B30	C30
NS-I M13F v2 NS-I M13R v2	11-529		✓	✓	✓	✓	✓	✓
NS-II M13F NS-II M13R	353-532		✓	✓	✓	✓	✓	✓
NS-II M13F A-NS-I-M13R	353-891		✓	✓	✓			
A-M-I-M13F A-M-I-M13R	1-1028	✓						
M-I M13F v2 M-I M13R v2	8-525		✓	✓	✓	✓	✓	✓
P30-M-I-F P30-M-I-R	426-801		✓	✓	✓			
P30-M-II-F P30-M-II-R	596-891		✓	✓	✓			
M-II M13F A-M-I-M13R	371-1028					✓	✓	✓
P30-M-II-F A-M-I-M13R	596-1028		✓	✓	✓			

protocol, used for amplification of segments using primers containing the M13 tags, consisted of an initial denaturation at 98°C for 30 seconds followed by five cycles of 98°C for 10 seconds, 50°C for 20 seconds, and 72°C for 30 seconds followed by 35 cycles of identical parameters, except for an increase in annealing temperature to 60°C, and a final elongation at 72°C for five minutes. The alternate protocol, used for amplification of segments not containing the M13 tags, consisted of an initial denaturation at 98°C for 30 seconds followed by 40 cycles of 98°C for 10 seconds, 60°C for 20 seconds, and 72°C for 30 seconds and a final elongation at 72°C for five minutes. All thermocycling was done using the Bio-Rad T100^{TM} Thermal Cycler (Cat #1861096). After PCR, amplified products were separated on a 2% agarose gel using gel

electrophoresis, excised from the gel, and purified using the Invitrogen PureLink™ Quick Gel Extraction Kit (Cat # K210012). Elution of the DNA was done in molecular grade water instead of elution buffer, however. Purified DNA was diluted to about 20 ng/µl using molecular grade water and sent to Eton Bioscience (5829 Oberlin Dr, San Diego, CA 92121) for sequencing. Sequencing data were compiled and assessed using the DNASTAR Lasergene and/or the Serial Cloner software.

RESULTS

Changes in Replicative Capacity

Plaque-to-plaque bottlenecked virus

To evaluate the replicative capacity of three viral lines (A, B, and C) after 30 plaque-to-plaque transfers, MDKC cells were infected at an MOI of 0.001 with passage 30 virus. Samples were taken every 12 hours post-infection out to 48 hours and characterized by routine plaque assay. Two of three viral lines exhibited lower viral titers at all hours post infection compared to the parental virus A/Victoria/361/2011 (H3N2) (Fig. 2). Both lines B and C had viral titers nearly two-log lower than that of the parental virus at 12 hours post infection. By 48 hours post infection, the viral titer differences of these two lines represented a roughly one-log decrease compared to that of the parental virus, with line C having slightly lower titers than that of line B. These decreased viral titers suggest the two viral lines experienced a drop in replicative capacity as a result of the serial passaging process. The remaining viral line, however, had viral titers similar to those of the parental virus, except at 12 hours post-infection, where the viral titer was roughly half a log lower than that of the parental virus. Based on these titers, this line did not appear to lose fitness after being subjected to repeated plaque-to-plaque bottlenecks (Fig. 2).

Differences in replicative capacity were also reflected in the phenotypic differences between the plaques produced by the serially bottlenecked viruses and those

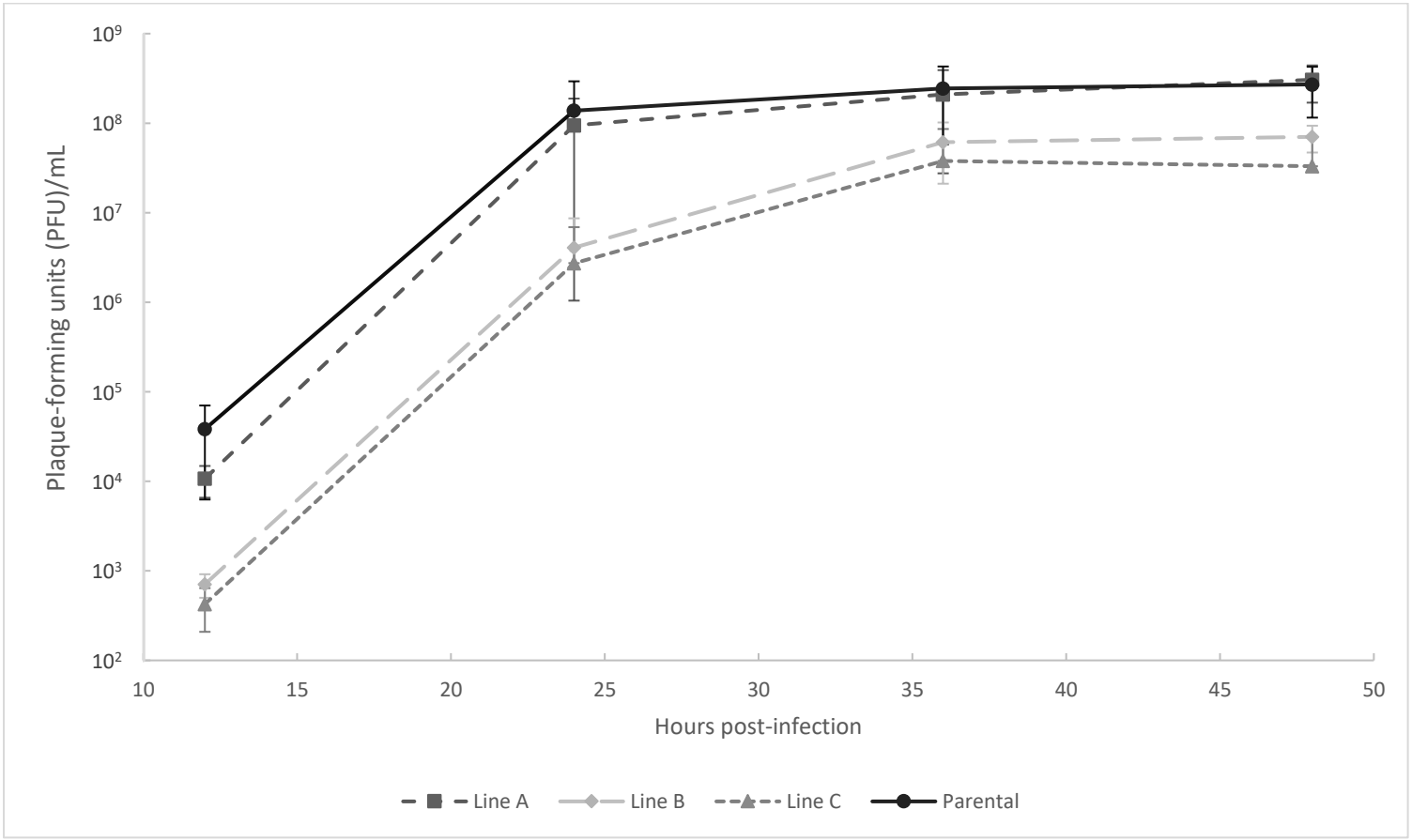

FIGURE 2 Multi-cycle growth curve of three independent viral lines (Line A/B/C) subjected to 30 plaque-to-plaque transfers, representing a bottleneck size of 1 PFU, compared to the parent virus A/Victoria/361/2011 (H3N2) (Parental). Flasks containing confluent MDCK cells were infected with bottlenecked viruses or the parental virus at an MOI of 0.001. Virus was harvested at the indicated hours post-infection and titered as plaque-forming units/mL (PFU/mL).

produced by the parental virus. Plaques produced by lines B and C, the viral lines exhibiting lower viral fitness compared to the parental virus, were both smaller and fainter than those produced by the parental virus (Fig. 3). Visual differences in the plaques are quite evident and support the conclusion that these viral lines have experienced a loss in fitness. Line A also produced plaques noticeably different in size than those produced by the parental virus, though plaque numbers matched those of the parental virus (Fig. 3). This phenotypic difference could suggest a change in replicative capacity that was not evident in the multi-cycle growth curve.

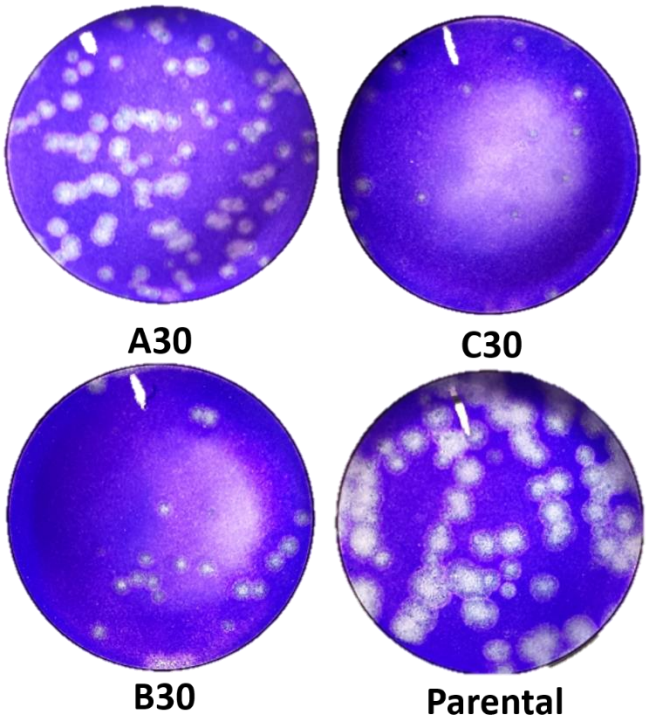

FIGURE 3 Plaque phenotypes of 1 PFU passaged viral lines compared to parental virus. Viral dilutions were prepared and inoculated onto MDKC cells within 12-well tissue culture plates. Following three days of growth, cells were fixed with formalin for a minimum of four hours and stained with crystal violet dye to visualize plaques.

To determine viral fitness over a single replicative cycle, viruses subjected to 30 plaque-to-plaque transfers were used to infect MDCK cells at an MOI of 5. The results of the multi-cycle growth curve were not mirrored in the single cycle growth curve (Fig. 4). All three serially passaged viruses showed lower PFU counts than the parental virus at the first time point, representing a roughly half-log drop in viral titer compared to the parental virus. However, by nine hours all three viruses had surpassed the parental virus, with line A surpassing the parental virus after just 6 hours (Fig. 4). Line A produced viral titers nearly a half-log greater than the parental virus at six hours, which increased to a one-log difference by twelve hours. Meanwhile, lines B and C had lower viral titers at six hours post infection, which changed into a roughly one-log increase over the parental virus by twelve hours post-infection. Like the multi-cycle growth curve, there also appeared to be a significant difference amongst the three bottlenecked viral lines. Line A had higher viral titers at all time points, most notably at six hours post-infection where viral titers were one-log higher than lines B and C, suggesting differences in replicative capacity that were evident in the multi-cycle curve as well. The visual nature of the growth curves themselves between line A and lines B and C also underline the differences between the viral lines. Viral titers for line A increased from three to nine hours post-infection before levelling out at twelve hours. Lines B and C had viral titers that remained steady between three and six hours post-infection, increased between six and nine hours, and remained steady once again between nine and twelve hours.

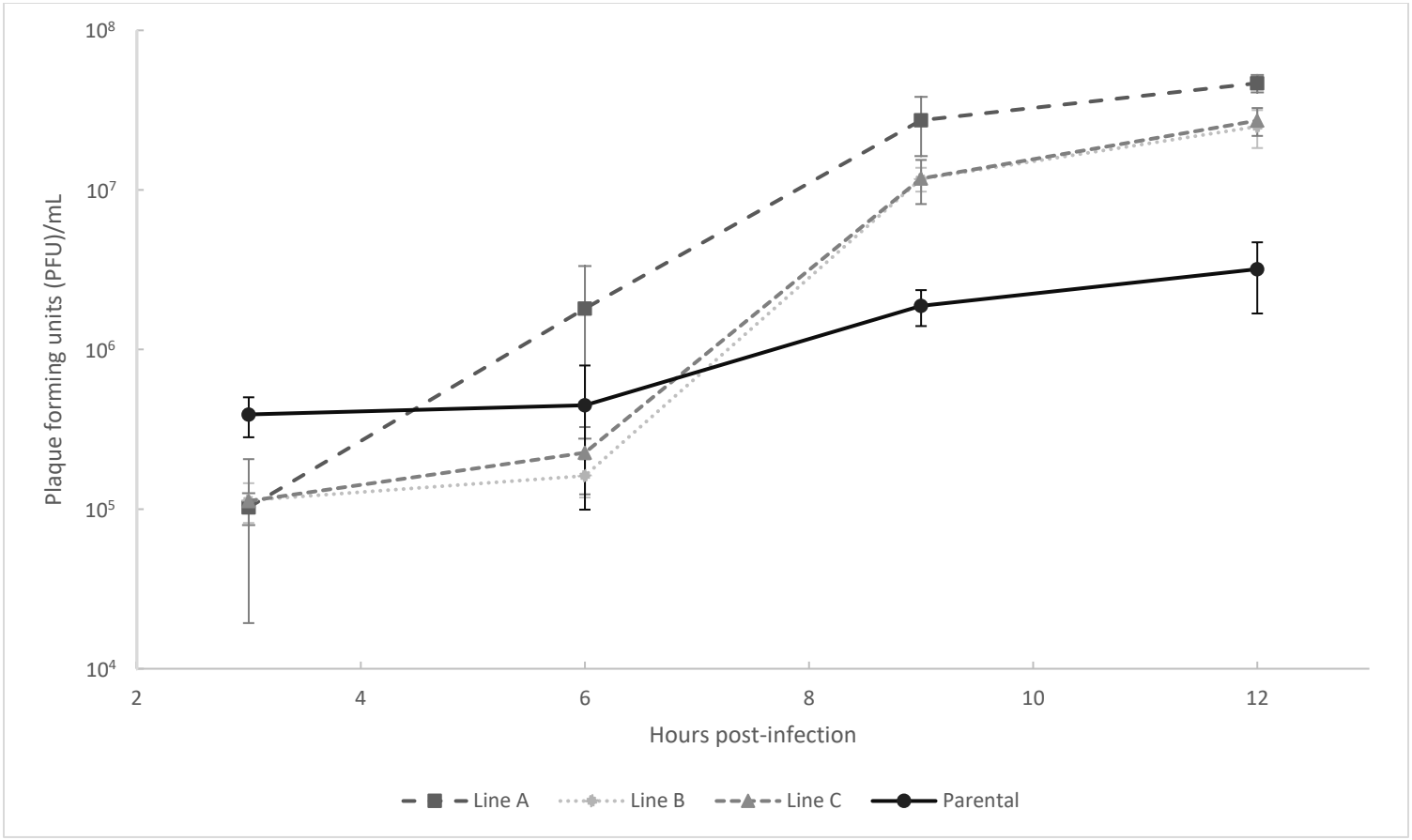

FIGURE 4 Single-cycle growth curve of three independent viral lines (Line A/B/C) subjected to 30 plaque-to-plaque transfers, representing a bottleneck size of 1 PFU, compared to the parent virus A/Victoria/361/2011 (H3N2) (Parental). Flasks containing confluent MDCK cells were infected with bottlenecked viruses or the parental virus at an MOI of 5. Virus was harvested at the indicated hours post-infection and titered as plaque-forming units/mL (PFU/mL).

1000 PFU bottlenecked virus

To evaluate the replicative capacity of viral lines after 30 serial passages at a bottleneck size of 1000 PFU, three serially passaged viral lines were used to infect MDKC cells at an MOI of 10^{-3}. After 30 passages at a bottleneck size of 1000 PFU, the serially passaged viruses exhibited increased titers compared to the parental virus, suggesting this bottleneck size was large enough to allow for an increase in replicative capacity as a result of the passaging process (Fig. 5). At all hours post-infection, the bottlenecked viral lines had higher viral titers than the parental virus. This increase was less than half a log at 12 hours, between the passaged viruses and the parental virus, but increased to a roughly half-log difference at 24, 36, and 48 hours post-infection. There were also no significant differences in replicative capacity, based on viral titer produced during the growth curve, amongst the three serially passaged lines. Not only were viral titers clustered between the three lines at all time points, but no one viral line had consistently higher titers across all time points (Fig. 5). This suggests the replicative capacity of the three serially bottlenecked viral lines to be roughly equivalent.

The phenotypes of the plaques produced by the serially passaged viruses, particularly when compared to the plaques produced by the parental virus, support the conclusion that the bottlenecked viruses improved their replicative capacity as a result of the passaging process. Plaques from bottlenecked viral lines A and B were visibly larger and more pronounced than those produced by the parental virus, suggesting a

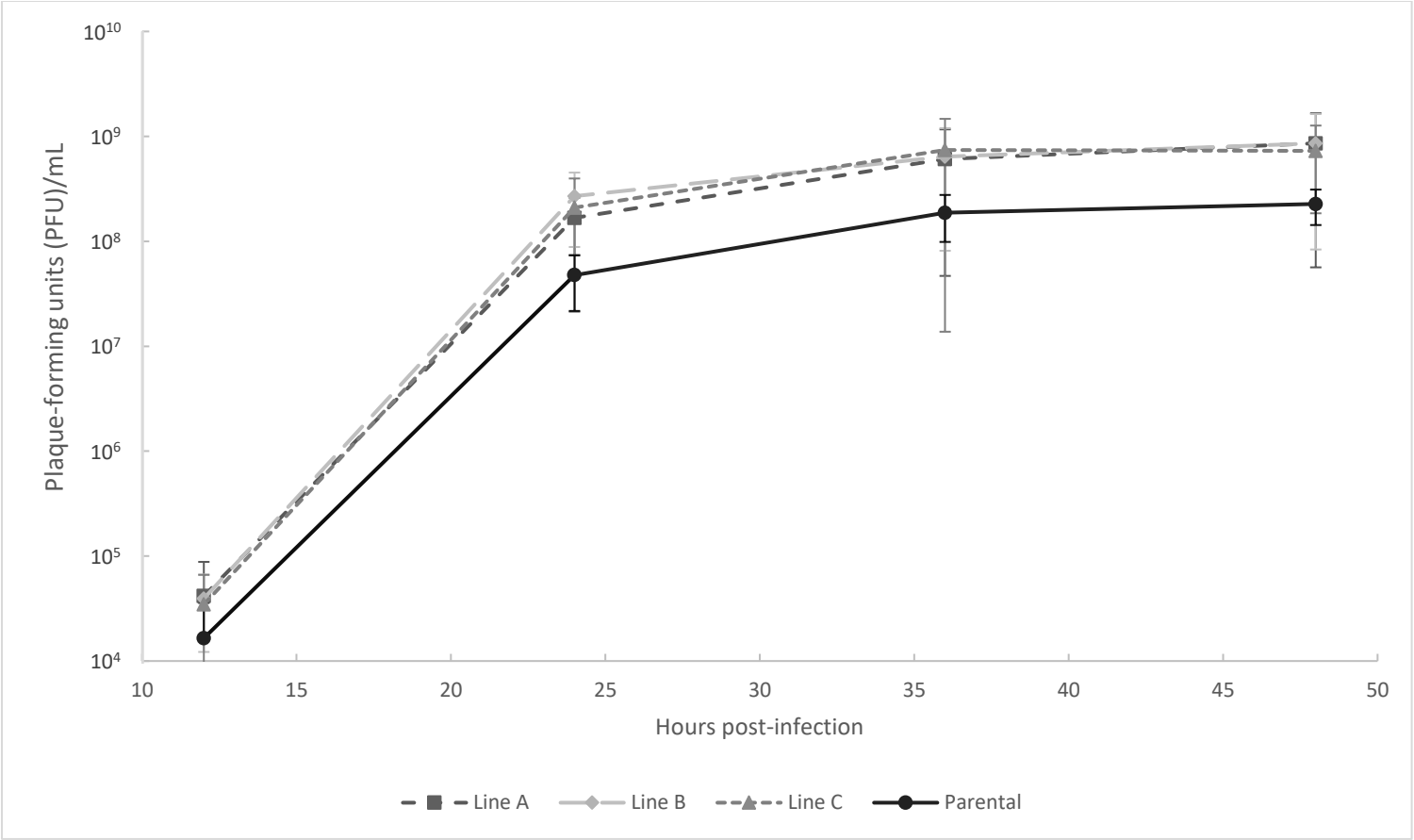

FIGURE 5 Multi-cycle growth curve of three independent viral lines (Line A, B, and C) subjected to thirty 1000 PFU bottlenecks compared to the parent virus A/Victoria/361/2011 (H3N2) (Parental). Flasks containing confluent MDCK cells were infected with bottlenecked viruses or the parental virus at an MOI of 0.001. Virus was harvested at the indicated hours post-infection and titered as plaque-forming units/mL (PFU/mL).

difference in replicative capacity (Fig. 6). Line C produced a number of plaques that were visibly similar to those produced by the parental virus, yet also produced some plaques with phenotypes closer to those produced by lines A and B. The smaller plaques produced by line C could be suggestive of a difference in replicative capacity from the other bottlenecked viral lines not evident in the multi-cycle growth curve (Fig. 6). It is also possible the phenotypic differences within the plaques formed by line C were

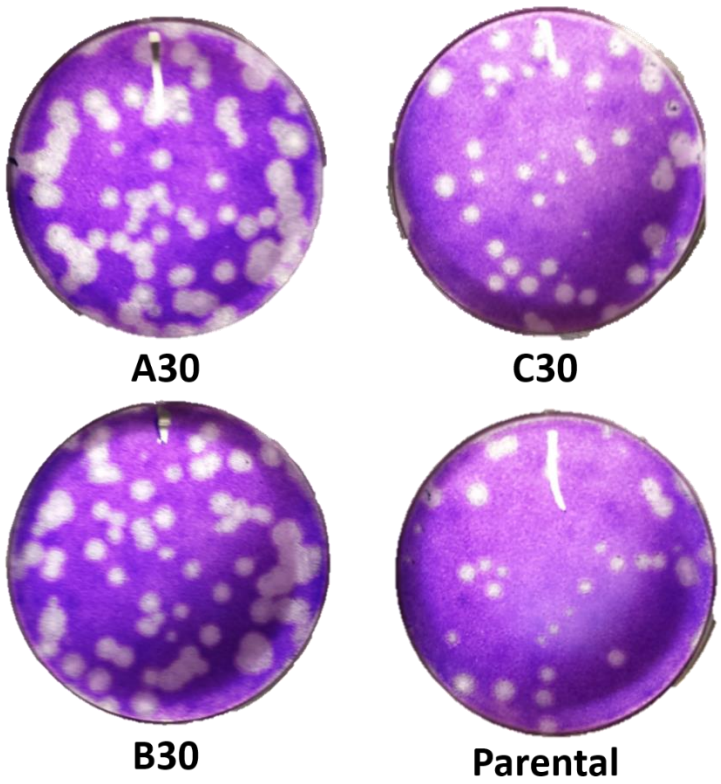

FIGURE 6 Plaque phenotypes of viral lines after thirty 1000 PFU passages (A30, B30, C30) compared to that of the parent virus A/Victoria/361/2011 (H3N2) (Parental). Viral dilutions were prepared and inoculated onto MDKC cells within 12-well tissue culture plates. Following three days of growth, cells were fixed with formalin for a minimum of four hours and stained with crystal violet dye to visualize plaques.

due to extraneous factors, such as agar concentration or cell monolayer thickness not related to the virus itself.

To determine viral fitness over a single replicative cycle, viruses subjected to thirty serial transfers at a size of 1000 PFU and then amplified once were used to infect MDCK cells at an MOI of 5. The single cycle growth curve supported the results of the multi-cycle growth curve and the difference in plaque phenotypes (Fig. 7). All three viral lines subjected to thirty 1000 PFU bottlenecks showed higher viral titers than the parental virus from six hours post-infection onwards, suggesting an improvement in replicative capacity as a result of the passaging process. Viral titers of the serially passaged viruses were nearly 1.5-log higher than those of the parental virus at six, nine, and twelve hours post-infection. In addition to the increased viral titers, the bottlenecked viruses exhibited a different growth curve than that seen in the parental virus. While the parental virus titers appeared to go through a lag phase before increasing after nine hours, the serially passaged viruses continually increased in count throughout the twelve hour growth curve (Fig. 7). This difference could also be indicative of a difference in replicative capacity. While the three viral lines were nearly identical at all hours post infection in the multi-cycle growth curve, there were minor discrepancies between the three viral lines in the single-cycle curve. In particular, both lines B and C produced higher viral titers at 9 and 12 hours post infection than line A, though this difference in titer was not nearly as pronounced as that seen between the bottlenecked viruses and the parental virus, which was nearly a two-log difference.

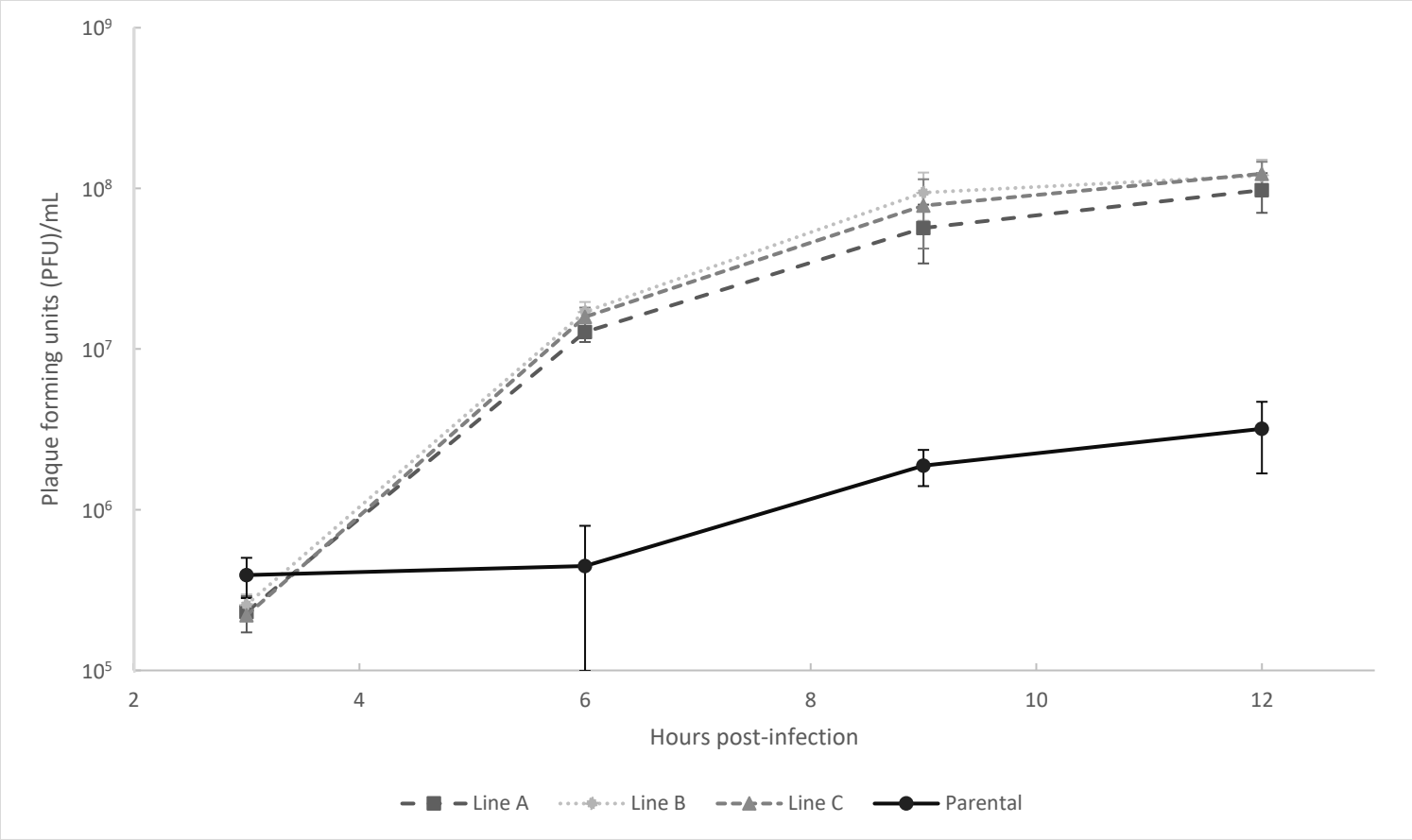

FIGURE 7 Single-cycle growth curve of three independent viral lines (Line A/B/C) subjected to thirty 1000 PFU bottlenecks compared to the parent virus A/Victoria/361/2011 (H3N2) (Parental). Flasks containing confluent MDCK cells were infected with bottlenecked viruses (Line A/B/C) or the parental virus at an MOI of 5. Virus was harvested at the indicated hours post-infection and titered as plaque-forming units/mL (PFU/mL).

Changes in Genomic Sequence

Plaque-to-plaque bottlenecked virus

After 30 serial plaque-to-plaque passages, ten mutations were identified in lines A and C while thirteen mutations were identified in line B, which were associated with the previously discussed changes in replicative capacity. In line A, 50% of present mutations were non-synonymous compared to 54% in line B and 70% in line C. Amongst the three viral lines, mutations were present in all eight genome segments, affecting nine different proteins. Line A and line C had mutations in five genome segments affecting five proteins whereas line B had mutations in seven genome segments altering eight proteins. All mutations, denoted by the genome affected, the base change caused by the mutation, and the accompanying amino acid substitution, if applicable, can be seen in Table 2. Nucleotide numbering is based off publicly available reference sequences for the parental virus, A/Victoria/361/2011 (H3N2), though sequence information is based on data collected from the parental virus used in these experiments (46). All three viral lines contained the mutations A545C and A544G, within the HA genome segment, which together contribute to the amino acid change Arg156Gln. In addition to the A544G and A545C mutations, there were another two mutations shared amongst all viral lines within the HA genome segment – C647A and A733C. Amidst the mutations across all three viral lines there were 16 transversions (54%) - mutations involving a switch from a purine (A/G) to a pyrimidine (C/T), or vice versa (Table 3). Of these sixteen transversions, five were in line A (50% of all mutations), eight were in line B (62%), and five in line C (50%). Three transversions in each viral line were synonymous mutations, with the remaining representing non-synonymous mutations.

TABLE 2 Genomic mutations identified in the plaque-to-plaque passaged viruses.

	Line A		Line B		Line C	
Segment	**Mutation**	**Amino Acid Change**	**Mutation**	**Amino Acid Change**	**Mutation**	**Amino Acid Change**
PB2	A37G	-[1]				
	C897A	-				
			A1464T	R479S		
	T1706C	V560A				
PB1			A397G	-		
					T781A	-
					G1217A	A401T
PA			A817C	-		
					A1661G	D547G
			A2166G	-		
HA	A544G	R156H	A544G	R156H	A544G	R156H
	A545C	-	A545C	-	A545C	-
	A544G & A545C	R156Q	A544G & A545C	R156Q	A544G & A545C	R156Q
	C647A	E190D	C647A	E190D	C647A	E190D
	A733C	Y219S	A733C	Y219S	A733C	Y219S
			G1299T	D408Y		
NA					A210G	E64G
			A265T	-		
	G968A	V317M			T958G	-
NP	A435T	-				
			A1401C	K452N		
M1			G349A	T108T		
					A546G	K174R
M2			C389T	-		
NS1	C253T	-				
Total	10		13		10	

1 Dashes indicate synonymous mutations, where no change in amino acid occurred

TABLE 3 Synonymous and non-synonymous mutations and transversions in plaque-to-plaque passaged viruses

	Line A	Line B	Line C
Synonymous	5 (50%)	6 (46%)	3 (30%)
Non-synonymous	5 (50%)	7 (54%)	7 (70%)
Transversions	5 (50%)	8 (62%)	5 (50%)

1000 PFU bottlenecked virus

After 30 serial passages at a bottleneck size of 1000 PFU, four mutations became fixed within the viral population of line A, six in line B, and five in line C. Between all three viral lines, mutations spanned five genome segments and affected five proteins. These mutations, their locations in the genome, base substitutions, and amino acid changes can be seen in Table 3. Nucleotide and amino acid numbering are based off reference sequencing information for the parental virus A/Victoria/361/2011 (H3N2) (46). Of the 15 total mutations amongst the three viral lines, 13 were non-synonymous, with the two synonymous mutations being located at nucleotides 1989 in the PB2 segment of line B and 1006 in the NA genome segment of line C. Additionally, each line had only two synonymous mutations present in non-HA genome segments. All remaining mutations for each of the three viral lines affected the HA protein. Amongst the three lines, there were five unique mutations within the HA genome segment. Two of these mutations were shared between lines B and C. Line A harbored two HA mutations not found in any of the other viral populations, C460A and T553A, while line B had only one - T655C. Two other mutations outside of the HA genome segment were also shared between viral lines. Both lines A and B contained the mutation G2245A within the PB2 genome segment and all three contained a mutation in the M segment, A546G. Amongst the mutations across all the three viral lines there were three transversions (20%) (Table 5). Two of these occurred in line A (50%), both in non-synonymous mutations, while one occurred in line B (17%), in the singular synonymous mutation present in that line. No transversions were present in line C.

TABLE 4 Genomic mutations identified in the 1000 PFU bottlenecked viruses.

	Line A		Line B		Line C	
Genome Segment	**Mutation**	**Amino Acid Change**	**Mutation**	**Amino Acid Change**	**Mutation**	**Amino Acid Change**
PB2			T1989G	-[1]		
	G2245A	D740N	G2245A	D740N		
PA					A107G	K29R
HA	C460A	T128N				
	T553A	F159Y				
			C572T	V165N	C572T	V165N
			G633A	V186I	G633A	V186I
			T655C	F193S		
NA					C1006T	-
M1	A546G	K174R	A546G	K174R	A546G	K174R
Total	4		6		5	

1 Dashes indicate synonymous mutations, where no change in amino acid occurred

TABLE 5 Synonymous and non-synonymous mutations and transversions in 1000 PFU passaged viruses.

	Line A	Line B	Line C
Synonymous	0 (0%)	1 (17%)	1 (20%)
Non-synonymous	4 (100%)	5 (83%)	4 (80%)
Transversions	2 (50%)	1 (17%)	0 (0%)

DISCUSSION

The primary goals of this research were to determine the effects of controlled artificial bottlenecks on the replicative capacity and evolutionary trajectory of the parental influenza virus A/Victoria/361/2011 (H3N2). This work has shown that at a bottleneck size of one virus, brought about through plaque-to-plaque passaging, reduced replicative capacity in two out of three viral lines was concomitant with a number of mutations throughout the viral genomes. Conversely, in viral lines passaged at a bottleneck size of 1000 PFU, replicative capacity was not only maintained, but actually increased. The mutations within these lines were primarily clustered in the HA genome segment, as opposed to being spread throughout the genome.

Fitness Effects of Plaque-to-Plaque Transfers and Associated Mutations

Based on previous work with other RNA viruses, it is no surprise that influenza also experiences a loss in fitness after repeated plaque-to-plaque transfers (39–41). It is interesting that only two of the three viral lines subjected to plaque-to-plaque bottlenecks exhibited a loss in fitness, but not unexpected. While it would be difficult for a virus to avoid the accumulation of deleterious mutations that would lead to a decrease in fitness when being repeatedly passaged at a bottleneck size of 1 PFU, it is not impossible, and serves to underscore the random nature of mutations in RNA viruses.

The mutations seen in the plaque-to-plaque viruses are suggestive of an absence of natural selection, which would be expected after such severe bottlenecks as those

produced through plaque-to-plaque transfers. Recent work determining the nature of influenza bottlenecks between humans found similar results, specifically that between-host bottlenecks were stringent enough to lead to the fixation of stochastic mutations and preventing positive selection of higher fitness variants (47). Even though high fitness mutations may appear within the population, the stringency and stochastic nature of the bottleneck impede the transmission of these high fitness clones, since those viruses passed on are based solely on their frequency in the original population. However, even in human infections, viruses containing beneficial mutations do not appear to reach frequencies above 2%, though antigenic drift still proceeds as a result of the sheer number of transmissions that occur during seasonal epidemics (47). Thus, in the presence of stringent bottlenecks, as seen in these plaque-to-plaque transfers, fitness fails to improve due to viruses containing beneficial mutations being present at such low frequencies. Why clones containing beneficial mutations don't reach frequencies high enough to promote their transmission remains an unanswered question. In the plaque-to-plaque model, it is possible the virus simply cannot replicate to high enough numbers for any variant to out populate lower fitness clones. One possibility could be that viral populations within individuals, and presumably within cell culture models, are dominated more by stochastic processes than positive selection.

The similar number of synonymous mutations and non-synonymous mutations in the plaque-to-plaque passaged viruses is also evidence of a lack of positive selection. In general, a higher rate of non-synonymous mutations (d_N) compared to synonymous mutations (d_S) at a specific codon is indicative of positive selection acting at that site within the genome. While d_N and d_S cannot be calculated for these viruses, primarily

because mutations within these viruses are not being tracked over time, the nearly equal number of synonymous and non-synonymous mutations is highly suggestive of an absence of positive selection acting on these viruses, which would be expected based on the stochastic nature of mutations within viral populations as referenced above.

The direct contribution of each individual mutation is hard to determine without using reverse genetics to investigate each mutation individually. However, since these mutations arose stochastically, their appearance within the genome and effect on the virus are random. Thus, most of these mutations aren't worth discussing further in terms of the evolutionary trajectory of the influenza virus, as no selective process improves their chances of becoming fixed within the viral populations.

Three notable mutations are A544G, A545C, and A647A within the HA genome segment of all three plaque-to-plaque passaged viruses. Reference data available from the Influenza Research Database for the parental virus A/Victoria/361/2011 (H3N2) shows the consensus sequence for this particular strain of the virus to have a GC at nucleotides 544 and 545 and a C at nucleotide 647. However, when the parental virus stock was sequenced, there were mutations at each of these positions within the genome – AA at 544/545 and A at 647. The chromatographic data at each of these locations showed the reference sequence present at low levels within the parental viral population. Their presence in the parental virus, albeit at low levels, suggests these mutations simply arose during the passaging process, and not as a result of the passaging process, as is the case for the rest of the mutations that were not present in the parental virus in any capacity. It is interesting, then, that these three mutations are the only three to be located at antigenic sites within the HA protein (48–50). In the absence of selective pressure, such as that

presented by human antibodies, it would be unexpected to find a significant number of mutations within antigenic sites. Thus, the fact that the three mutations present in the antigenic site of the HA protein likely did not arise as a result of stochastic processes, as is probably the case with the rest of the mutations, but was present in some capacity within the parental population already supports the idea that strict bottlenecks limit selection and promote stochastic mutation.

Single Cycle Discrepancies in Plaque-to-Plaque Passaged Virus

What is quite interesting, and counterintuitive, is the apparent improvement in replicative capacity of all three viral lines (relative to the parental virus) in the single-cycle growth curve versus the multi-cycle growth curve. There are a couple of potential explanations for this apparent discrepancy between the two curves, however.

First, separate virus stocks were used for each growth curve and could account for the differences. For the multi-cycle growth curve, virus from the passage 30 plaques were picked and then briefly amplified for 24 hours to provide enough viral material to use in subsequent experiments, including sequencing. However, there were not enough viruses present in this primary amplified stock to use in the single-cycle growth curve, which required enough virus to reach an MOI of 5 for the infection. To overcome this, the primary amplified virus was amplified again for 36 hours before being harvested and used for the single-cycle growth curve. It is possible that the viral lines were able to improve in fitness, as has been observed in large-scale passages of low fitness viral populations.

Alternatively, the consistent low titers at all time points for the parental virus could be due to defective interfering particles (DIPs). These particles are virions containing all the necessary surface proteins for attachment and entry with a genome that is replication deficient due to deletions or other genetic mutations(51). Influenza DIPs most frequently contain deletions in segments 1-3, which code for the subunits of the viral RNA-dependent RNA polymerase, thus rendering them unable to replicate on their own. These DIPs can only replicate in the presence of functional, infectious virus. Interference from defective viruses can occur when replication of defective genome segments exceeds that of functional segments (51). It is possible that at high MOIs the ratio of DIPs to functional virus is high enough to produce this interference, but low enough at low MOIs (like in the multi-cycle growth curve) that the growth of the parental virus is not affected. If this were the case, it would explain the differences between the two growth curves for the parental virus. Such particles would be absent for the passaged viral lines since the passaging process would naturally eliminate defective particles, due to the requirement of co-infection for the replication of DIPs.

Fitness Effects of Passaging at a Bottleneck Size of 1000 PFU

The increase in replicative capacity of the viral lines after repeated bottlenecks at a size of 1000 PFU clearly suggest that this bottleneck size is at or above the threshold required for maintenance or improvement of fitness in the influenza A virus. Beneficial mutations are more likely to become fixed in large populations than in small populations (52). Such a difference is precisely the result observed with this research when comparing the fitness changes in the plaque-to-plaque viruses, where population sizes are limited, and those passaged at 1000 PFU, where larger populations are produced during infection.

Furthermore, as was previously discussed, viruses harboring mutations with higher fitness effects often fail to reach frequencies above 2% in viral populations within human hosts. This is presumably also the case within viral populations produced in cell culture. However, while transmission between humans produce bottlenecks allowing only one to two viral genomes to found a new population, this research dealt with the transmission of 1000 viral genomes (47). Clearly a bottleneck of this size was more than enough to allow the transmission of those viruses containing beneficial mutations, allowing for the improvement of fitness over time.

Mutations Associated with Passaging at a Bottleneck Size of 1000 PFU

Contrary to what was seen in the plaque-to-plaque viruses, the mutations present in the viruses passaged at the 1000 PFU bottleneck are highly suggestive of strong positive selection. As was discussed previously, though rates of d_N and d_S cannot be calculated, the ratio of non-synonymous to synonymous mutations can be suggestive of positive selection. The majority of the mutations in these viral lines are non-synonymous, which would be expected if a selective pressure was acting on them.

Additionally, many of the mutations in all three of these viral lines are present in the HA genome segment, a well-known target of antigenic change in influenza A. All five of the unique HA mutations amongst the three viral lines were located in or around antigenic sites of the H3 protein, which is strange considering there is no selective pressure from antibodies in MDCK cells (48–50). It is clear that these mutations must still provide the virus some advantage in tissue culture separate from that afforded viruses adapting to antibody pressure. These mutations were likely the result of further adaptation to MDCK cells, as antigenic variation during adaption to new host species or

new cell lines (as would be the case in laboratory work) is a well-known process in influenza A (53).

Several of the residues at which mutations are present in the passaged viruses are associated with changes in virulence. Mutations at residues 128, 165, 186 and 193 in influenza A viruses with an H3 protein are associated with changes in viral binding to α2-6 glycans, altering receptor-binding specificity (54–56). Alterations in receptor-binding specificity, leading to an increase in binding to certain oligosaccharides or broadening the range of recognized oligosaccharides, can improve virulence (55). Such changes may be what occurred in these viruses - further adaptation to MDCK cells led to improvement in receptor-binding specificity. In addition to changes in viral binding, previous work has determined that a T128N mutation results in a loss of a potential glycosylation site for the H3 protein (57). An increased number of glycosylation sites within the H3 protein of H3N2 viruses has been associated with decreased virulence in these viruses (58). The loss of a glycosylation site could lead to an increase in virulence, and such a change in the passaged viruses could contribute to the observed increase in replicative capacity.

The M1 matrix protein of influenza A has a number of important roles in the viral replicative cycle. Export of newly replicated ribonucleoprotein from the host cell nucleus and budding of new viral particles is dependent on the M1 protein (59, 60). The mutation K174R present in all three viral lines is located within the C-terminal domain of the M1 protein, which covers residues 165-252. The function of the C-terminal domain is still unclear, though there is some evidence that this domain interacts with the host cell, is important in virus assembly, and contributes to oligomerization of the M1 protein.

Alterations to any of these functions in a beneficial manner could certainly impact replicative capacity in the manner observed through the growth curves.

Significance of Transversions

It has been widely observed that transitions, purine to purine or pyrimidine to pyrimidine changes, occur more frequently, despite there being more opportunities for a transversion to occur, known as the transition bias (61). There are two main hypotheses for why this bias exists. The mutational hypothesis suggests polymerases are more prone to making transition mutations than they are transversion mutations, while the selectional hypothesis posits that natural selection favors transitions over transversions, supposing that transitions are less biochemically severe and lead to less chemical changes in amino acids (61). With these hypotheses in mind, it is worth questioning if either is in play when considering the differences in transversions between the plaque-to-plaque passaged viruses and those passaged at a bottleneck size of 1000 PFU. While the sample size is very small for the 1000 PFU viruses, there are fewer transversions when compared to the plaque-to-plaque viruses. Previous work on influenza A found that transitions are less detrimental than transversions (62). If this is the case, and the selectional hypothesis were true, it would be expected that transversions would be selected against within viral populations, if positive selection were allowed to act, which appears to be what is occurring in these serially passaged viral lines. In the plaque-to-plaque viruses, selection cannot take place, so transversions become more numerous, but in the 1000 PFU viral lines, positive selection is acting and reducing the number of transversions, favoring transitions instead.

Through the serial passaging of the influenza A virus in artificial bottlenecks of 1 and 1000 viruses, we have shown that bottleneck size drives the direction of evolution *in vitro*. At small bottleneck sizes, the stochastic nature of mutations is the driving force behind evolution of the virus, leading to fitness decreases over time. However, at larger bottleneck sizes evolution is driven by positive selection of more beneficial mutations, resulting in fitness increases. Mutations associated with these fitness changes reflect the driving force behind the evolution of the virus – at small bottleneck sizes, mutations are spread throughout the genome reflecting their stochastic nature, whereas at larger bottleneck sizes, mutations are clustered in sites with well-documented effects on virulence, which would be expected if positive selection were acting.

Future work on influenza and genetic bottlenecks should be focused on identifying the bottleneck size where the tradeoff between fitness loss and fitness gain occurs. Much of the previous research on bottlenecks and influenza identified new infections as being founded by fewer than 100 viral particles. Therefore, it is not surprising the virus maintains and even improves fitness when passaging 1000 particles. Presumably at some point between 1 and 1000 viruses, likely closer to that 10-100 particle number found in respiratory droplets, there occurs a shift where bottlenecks no longer allow for the maintenance or improvement of fitness. However, the error-prone nature of influenza genome replication combined with large population sizes achieved during infection may compensate for the initial loss of viral fitness in the viral population that establishes infection following a bottleneck event (e.g. through reversion of mutated sites). This research and future work on bottlenecks is significant for both understanding the evolution of currently circulating influenza strains that cause seasonal epidemics and

the emergence of avian influenza pandemic strains from animal reservoirs. Improving our knowledge of the effect of bottlenecks that occur during transmission of influenza can help our understanding of their impact on seasonal epidemics and illuminate their role as a roadblock that must be overcome by avian influenza viruses to reach pandemic potential.

REFERENCES

1. Bouvier NM, Palese P. 2008. The Biology of Influenza Viruses. Vaccine 26:D49–D53.

2. Taubenberger JK, Morens DM. 2008. The Pathology of Influenza Virus Infections. Annu Rev Pathol 3:499–522.

3. Zambon MC. 1999. Epidemiology and pathogenesis of influenza. J Antimicrob Chemother 44:3–9.

4. Types of Influenza Viruses | Seasonal Influenza (Flu) | CDC.

5. Couch RB. 1996. Orthomyxoviruses, p. . *In* Baron, S (ed.), Medical Microbiology, 4th ed. University of Texas Medical Branch at Galveston, Galveston (TX).

6. Nicholson KG, Wood JM, Zambon M. 2003. Influenza. The Lancet 362:1733–1745.

7. Rello J, Pop-Vicas A. 2009. Clinical review: Primary influenza viral pneumonia. Crit Care 13:235.

8. Key Facts About Influenza (Flu) | Seasonal Influenza (Flu) | CDC.

9. Molinari N-AM, Ortega-Sanchez IR, Messonnier ML, Thompson WW, Wortley PM, Weintraub E, Bridges CB. 2007. The annual impact of seasonal influenza in the US: Measuring disease burden and costs (Economic Ref). Vaccine 25:5086–5096.

10. Grohskopf LA. 2016. Prevention and Control of Seasonal Influenza with Vaccines. MMWR Recomm Rep 65.

11. Thompson WW, Shay DK, Weintraub E, Brammer L, Cox N, Anderson LJ, Fukuda K. 2003. Mortality Associated With Influenza and Respiratory Syncytial Virus in the United States. JAMA 289:179–186.

12. Thompson WW, Shay DK, Weintraub E, Brammer L, Bridges CB, Cox NJ, Fukuda K. 2004. Influenza-Associated Hospitalizations in the United States. JAMA 292:1333–1340.

13. Morens JKT and DM. 1918 Influenza: the Mother of All Pandemics - Volume 12, Number 1—January 2006 - Emerging Infectious Disease journal - CDC.

14. Shrestha SS, Swerdlow DL, Borse RH, Prabhu VS, Finelli L, Atkins CY, Owusu-Edusei K, Bell B, Mead PS, Biggerstaff M, Brammer L, Davidson H, Jernigan D, Jhung MA, Kamimoto LA, Merlin TL, Nowell M, Redd SC, Reed C, Schuchat A, Meltzer MI. 2011. Estimating the Burden of 2009 Pandemic Influenza A (H1N1) in the United States (April 2009–April 2010). Clin Infect Dis 52:S75–S82.

15. Tewawong N, Prachayangprecha S, Vichiwattana P, Korkong S, Klinfueng S, Vongpunsawad S, Thongmee T, Theamboonlers A, Poovorawan Y. 2015. Assessing Antigenic Drift of Seasonal Influenza A(H3N2) and A(H1N1)pdm09 Viruses. PLOS ONE 10:e0139958.

16. Treanor J. 2004. Influenza Vaccine — Outmaneuvering Antigenic Shift and Drift. N Engl J Med 350:218–220.

17. Hay AJ, Gregory V, Douglas AR, Lin YP. 2001. The evolution of human influenza viruses. Philos Trans R Soc Lond Ser B 356:1861–1870.

18. Samji T. 2009. Influenza A: Understanding the Viral Life Cycle. Yale J Biol Med 82:153–159.

19. Varki A. 2008. Sialic acids in human health and disease. Trends Mol Med 14:351–360.

20. Boerlijst MC, Bonhoeffer S, Nowak MA. 1996. Viral Quasi-Species and Recombination. Proc Biol Sci 263:1577–1584.

21. Lauring AS, Andino R. 2010. Quasispecies Theory and the Behavior of RNA Viruses. PLOS Pathog 6:e1001005.

22. Crotty S, Cameron CE, Andino R. 2001. RNA virus error catastrophe: Direct molecular test by using ribavirin. Proc Natl Acad Sci 98:6895–6900.

23. Van den Hoecke S, Verhelst J, Vuylsteke M, Saelens X. 2015. Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing. BMC Genomics 16:79.

24. Andino R, Domingo E. 2015. Viral quasispecies. Virology 479–480:46–51.

25. Li H, Roossinck MJ. 2004. Genetic Bottlenecks Reduce Population Variation in an Experimental RNA Virus Population. J Virol 78:10582–10587.

26. Brankston G, Gitterman L, Hirji Z, Lemieux C, Gardam M. 2007. Transmission of influenza A in human beings. Lancet Infect Dis 7:257–265.

27. Catching the Flu: NIOSH Research on Airborne Influenza Transmission | NIOSH Science Blog | Blogs | CDC.

28. Gustin KM, Katz JM, Tumpey TM, Maines TR. 2013. Comparison of the Levels of Infectious Virus in Respirable Aerosols Exhaled by Ferrets Infected with Influenza Viruses Exhibiting Diverse Transmissibility Phenotypes. J Virol 87:7864–7873.

29. Milton DK, Fabian MP, Cowling BJ, Grantham ML, McDevitt JJ. 2013. Influenza Virus Aerosols in Human Exhaled Breath: Particle Size, Culturability, and Effect of Surgical Masks. PLoS Pathog 9.

30. Lindsley WG, Blachere FM, Thewlis RE, Vishnu A, Davis KA, Cao G, Palmer JE, Clark KE, Fisher MA, Khakoo R, Beezhold DH. 2010. Measurements of Airborne Influenza Virus in Aerosol Particles from Human Coughs. PLoS ONE 5.

31. Moya A, Elena SF, Bracho A, Miralles R, Barrio E. 2000. The evolution of RNA viruses: A population genetics view. Proc Natl Acad Sci 97:6967–6973.

32. Schneider WL, Roossinck MJ. 2001. Genetic Diversity in RNA Virus Quasispecies Is Controlled by Host-Virus Interactions. J Virol 75:6566–6571.

33. Varble A, Albrecht RA, Backes S, Crumiller M, Bouvier NM, Sachs D, García-Sastre A, tenOever BR. 2014. Influenza A Virus Transmission Bottlenecks Are Defined by Infection Route and Recipient Host. Cell Host Microbe 16:691–700.

34. Wilker PR, Dinis JM, Starrett G, Imai M, Hatta M, Nelson CW, O'Connor DH, Hughes AL, Neumann G, Kawaoka Y, Friedrich TC. 2013. Selection on haemagglutinin imposes a bottleneck during mammalian transmission of reassortant H5N1 influenza viruses. Nat Commun 4:ncomms3636.

35. Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson JA, Ping L-H, Swanstrom R, Tomaras GD, Blattner WA, Goepfert PA, Kilby JM, Saag MS, Delwart EL, Busch MP, Cohen MS, Montefiori DC, Haynes BF, Gaschen B, Athreya GS, Lee HY, Wood N, Seoighe C, Perelson AS, Bhattacharya T, Korber BT, Hahn BH, Shaw GM. 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc Natl Acad Sci 105:7552–7557.

36. Crow JF. 2005. Hermann Joseph Muller, Evolutionist. Nat Rev Genet 6:941–945.

37. Butcher D. 1995. Muller's Ratchet, Epistasis and Mutation Effects. Genetics 141:431–437.

38. Gordo I, Navarro A, Charlesworth B. 2002. Muller's ratchet and the pattern of variation at a neutral locus. Genetics 161:835–848.

39. Escarmís C, Dávila M, Charpentier N, Bracho A, Moya A, Domingo E. 1996. Genetic Lesions Associated with Muller's Ratchet in an RNA Virus. J Mol Biol 264:255–267.

40. Clarke DK, Duarte EA, Moya A, Elena SF, Domingo E, Holland J. 1993. Genetic bottlenecks and population passages cause profound fitness differences in RNA viruses. J Virol 67:222–228.

41. Chao L. 1990. Fitness of RNA virus decreased by Muller's ratchet. Nature 348:454–455.

42. Novella IS, Duarte EA, Elena SF, Moya A, Domingo E, Holland JJ. 1995. Exponential increases of RNA virus fitness during large population transmissions. Proc Natl Acad Sci U S A 92:5841–5844.

43. Novella IS, Elena SF, Moya A, Domingo E, Holland JJ. 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. J Virol 69:2869–2872.

44. Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, Zhong G, Hanson A, Katsura H, Watanabe S, Li C, Kawakami E, Yamada S, Kiso M, Suzuki Y, Maher EA, Neumann G, Kawaoka Y. 2012. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. Nature 486:420–428.

45. Herfst S, Schrauwen EJA, Linster M, Chutinimitkul S, de Wit E, Munster VJ, Sorrell EM, Bestebroer TM, Burke DF, Smith DJ, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM. 2012. Airborne Transmission of Influenza A/H5N1 Virus Between Ferrets. Science 336:1534–1541.

46. Influenza Research Database - Nucleotide Sequence Search Results.

47. McCrone JT, Woods RJ, Martin ET, Malosh RE, Monto AS, Lauring AS. 2018. Stochastic processes constrain the within and between host evolution of influenza virus. eLife 7:e35962.

48. Wiley DC, Wilson IA, Skehel JJ. 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature 289:373–378.

49. Lednicky J, Iovine N, Brew J. Hemagglutinin Gene Clade 3C.2a Influenza A(H3N2) Viruses, Alachua County, Florida, USA, 2014–15 - Volume 22, Number 1—January 2016 - Emerging Infectious Disease journal - CDC.

50. Huang J-W, Yang J-M. 2011. Changed epitopes drive the antigenic drift for influenza A (H3N2) viruses. BMC Bioinformatics 12:S31.

51. Scott PD, Meng B, Marriott AC, Easton AJ, Dimmock NJ. 2011. Defective interfering influenza virus confers only short-lived protection against influenza virus disease: Evidence for a role for adaptive immunity in DI virus-mediated protection in vivo. Vaccine 29:6584–6591.

52. Foll M, Poh Y-P, Renzette N, Ferrer-Admetlla A, Bank C, Shim H, Malaspinas A-S, Ewing G, Liu P, Wegmann D, Caffrey DR, Zeldovich KB, Bolon DN, Wang JP, Kowalik TF, Schiffer CA, Finberg RW, Jensen JD. 2014. Influenza Virus Drug Resistance: A Time-Sampled Population Genetics Perspective. PLOS Genet 10:e1004185.

53. Archetti I, Horsfall FL. 1950. Persistant antigenic variation of influenza A viruses after incomplete neutralization in ovo with heterologous immune serum. J Exp Med 92:441–462.

54. Burke DF, Smith DJ. 2014. A Recommended Numbering Scheme for Influenza A HA Subtypes. PLOS ONE 9:e112302.

55. Mochalova L, Gambaryan A, Romanova J, Tuzikov A, Chinarev A, Katinger D, Katinger H, Egorov A, Bovin N. 2003. Receptor-binding properties of modern human influenza viruses primarily isolated in Vero and MDCK cells and chicken embryonated eggs. Virology 313:473–480.

56. Nobusawa E, Ishihara H, Morishita T, Sato K, Nakajima K. 2000. Change in Receptor-Binding Specificity of Recent Human Influenza A Viruses (H3N2): A Single Amino Acid Change in Hemagglutinin Altered Its Recognition of Sialyloligosaccharides. Virology 278:587–596.

57. Eshaghi A, Shalhoub S, Rosenfeld P, Li A, Higgins RR, Stogios PJ, Savchenko A, Bastien N, Li Y, Rotstein C, Gubbay JB. 2014. Multiple Influenza A (H3N2) Mutations Conferring Resistance to Neuraminidase Inhibitors in a Bone Marrow Transplant Recipient. Antimicrob Agents Chemother 58:7188–7197.

58. Alymova IV, York IA, Air GM, Cipollo JF, Gulati S, Baranovich T, Kumar A, Zeng H, Gansebom S, McCullers JA. 2016. Glycosylation changes in the globular head of H3N2 influenza hemagglutinin modulate receptor binding without affecting virus virulence. Sci Rep 6:36216.

59. Martin K, Heleniust A. 1991. Nuclear transport of influenza virus ribonucleoproteins: The viral matrix protein (M1) promotes export and inhibits import. Cell 67:117–130.

60. Gómez-Puertas P, Albo C, Pérez-Pastrana E, Vivo A, Portela A. 2000. Influenza Virus Matrix Protein Is the Major Driving Force in Virus Budding. J Virol 74:11538–11547.

61. Stoltzfus A, Norris RW. 2016. On the Causes of Evolutionary Transition:Transversion Bias. Mol Biol Evol 33:595–602.

62. Lyons DM, Lauring AS. 2017. Evidence for the Selective Basis of Transition-to-Transversion Substitution Bias in Two RNA Viruses. Mol Biol Evol 34:3205–3215.

APPENDIX A

PRIMER SEQUENCES

TABLE 6 Sequences of primers used in this study. Underlined sequence denotes the M13 Forward tag. Sequence in bold denotes the M13 Reverse tag.

Primer	Sequence (5' → 3')
A-Uni12F	5'-AGCRAAAGCAGG-3'
M13F	5'-GTAAAACGACGGCCAG-3'
M13R	5'-**CAGGAAACAGCTATGAC**-3'
A-PB2-I-M13F	5'-TGTAAAACGACGGCCAGTCTCGAGCAAAAGCAGGTCAA-3'
A-PB2-I-M13R	5'-**CAGGAAACAGCTATGAC**CGATGCTARTGGRTCTGCTG-3'
A-PB2-II-M13F	5'-TGTAAAACGACGGCCAGTGGGARCARATGTACACTC-3'
A-PB2-II-M13R	5'-**CAGGAAACAGCTATGAC**CYTGYGTTTCACTGAC-3'
A-PB2-III-M13F	5'-TGTAAAACGACGGCCAGTGCAAAATGGGWGTRGATG-3'
A-PB2-III-M13R	5'-**CAGGAAACAGCTATGAC**CTAGAAACAAGGTCGTT-3'
A-PB1-I-M13F	5'-TGTAAAACGACGGCCAGTAGCAAAAGCAGGCAAACCAT-3'
A-PB1-I-M13R	5'-**CAGGAAACAGCTATGAC**CTGTTCAAGCTTTTCRCAWATGC-3'
A-PB1-II-M13F	5'-TGTAAAACGACGGCCAGTCRATGACCAAAGATGCWGA-3'
A-PB1-II-M13R	5'-**CAGGAAACAGCTATGAC**CAAGGTCATTGTTTATCATRTTG-3'
A-PB1-III-M13F	5'-TGTAAAACGACGGCCAGTGTTGCYAATTTYAGCATGGAG-3'
A-PB1-III-M13R	5'-**CAGGAAACAGCTATGAC**CAGTAGAAACAAGGCATTT-3'
A-PA-I-M13F	5'-TGTAAAACGACGGCCAGTAGCRAAAGCAGGTACTGAT-3'
A-PA-I-M13R	5'-**CAGGAAACAGCTATGAC**CGGYTCTTTCCAKCCAAAG-3'
A-PA-II-M13F	5'-TGTAAAACGACGGCCAGTCMAARTTCCTSCTGATG-3'
A-PA-II-M13R	5'-**CAGGAAACAGCTATGAC**CTCMAGTCTYGGGTCTGTGAG-3'
A-PA-III-M13F	5'-TGTAAAACGACGGCCAGTGCATCCTGTGCAGCMATGGA-3'
A-PA-III-M13R	5'-**CAGGAAACAGCTATGAC**CAGTAGAAACAAGGTACCTTTT-3'
A-NP-I-M13F	5'-TGTAAAACGACGGCCAGTCAGGGTWRATAATCACTCAMTG-3'
A-NP-I-M13R	5'-**CAGGAAACAGCTATGAC**CTGRCTCTTGTGWGCTGG-3'
A-NP-II-M13F	5'-TGTAAAACGACGGCCAGTYTGAGRGGRTCAGTTGC-3'
A-NP-II-M13R	5'-**CAGGAAACAGCTATGAC**CAGTAGAAACAAGGGTATTTTTC-3'
A-M-I-M13F	5'-TGTAAAACGACGGCCAGTAGCAAAAGCAAAAGCAGGTAG-3'
A-M-I-M13R	5'-**CAGGAAACAGCTATGAC**CAGTAGAAACAAGGTAG-3'
A-NS-I-M13F	5'-TGTAAAACGACGGCCAGTAGCAAAAGCAGGGTGACAAAGACA-3'
A-NS-I-M13R	5'-**CAGGAAACAGCTATGAC**CAGTAGAAACAAGGGTGTTTTTTAT-3'
A-HA-I-M13F	5'-TGTAAAACGACGGCCAGTAAAGCAGGGGATAATTCTA-3'
A-HA-I-M13R	5'-**CAGGAAACAGCTATGAC**CATTGCTGCTTGAGTGCTT-3'
A-HA-II-M13F	5'-TGTAAAACGACGGCCAGTGGTTACTTCAAAATAC-3'
A-HA-II-M13R	5'-**CAGGAAACAGCTATGAC**CAGTAGAAACAAGGGTGTTTT-3'
A-NA-I-M13F	5'-TGTAAAACGACGGCCAGTAGCAAAAGCAGGAGT-3'
A-NA-I-M13R	5'-**CAGGAAACAGCTATGAC**CCGACATGCTGAGCACTYCCTGAC-3'
A-NA-II-M13F	5'-TGTAAAACGACGGCCAGTGAACTTGTRCAGTRGTAATG-3'
A-NA-II-M13R	5'-**CAGGAAACAGCTATGAC**CAGTAGAAACAAGGAG-3'

TABLE 6 (Continued)

Primer	Sequence (5' → 3')
NA-I M13F v2	5'-GTAAAACGACGGCCAGGCAGGAGTAAAGATGAATCC-3'
NA-I M13R v2	5'-**CAGGAAACAGCTATGAC**CCCTGACAATGTGCTAGTATGAAG-3'
NS-I M13F v2	5'-GTAAAACGACGGCCAGGGGTGACAAAGACATAATGGATTCC-3'
NS-I M13R v2	5'-**CAGGAAACAGCTATGAC**CCCTGGAAAAGAAGGCAATGGTGAG-3'
NS-II M13F v2	5'-GTAAAACGACGGCCAGGAAAGTGGAAGGACCTCTTTGC-3'
NS-II M13R v2	5'-**CAGGAAACAGCTATGAC**CCCTGTTCTACTTCAAACAGC-3'
M-I M13F v2	5'-GTAAAACGACGGCCAGGCAGGTAGATATTGAAAGATGAGC-3'
M-I M13R v2	5'-**CAGGAAACAGCTATGAC**CGCCACCATCTGCCTATGAGACC-3'
M-II M13F v2	5'-GTAAAACGACGGCCAGGCTCTCAGTTATTCTGCTGGTGC-3'
P30-PB1-II-R	5'-CCCATCCCACCAGTATGTTGTCTTGG-3'
P30-PB1-III-F	5'-GGCACAGCATCATTGAGCCCTGG-3'
P30-PB1-III-R	5'-GGTCCCCCATCTGATACCAATAGTCC-3'
P30-PA-I-F	5'-GCAGTAGCGAAAGCAGGTACTG-3'
P30-PA-I-R	5'-GGCAGGAGAAGTTCGGTGGG-3'
P30-PA-II-F	5'-GGCCTCTGGGATTCCTTTCG-3'
P30-PA-II-R	5'-GCTCAATTGGGGCTACGTCC-3'
P30-PA-III-F	5'-GGGCTCTTGGTGAAAACATGGC-3'
P30-PA-III-R	5'-CGGCTTCAATCATGCTCTCGATCTGC-3'
P30-PA-IV-F	5'-GTGCCATAGGCCAAATTTCAAGACCG-3'
P30-PA-IV-R	5'-GGACAGTACGGATAACAAATAGCAGC-3'
P30-PB2-I-F	5'-AGCAAAAGCAGGTCAATTATATTCAGTATGG-3'
P30-PB2-I-R	5'-GGTGGAGGAGTGAGGAATGACG-3'
P30-PB2-II-F	5'-CCCTTGATGGTCGCATACATGTTAGAG-3'
P30-PB2-II-R	5'-GGGGTTCAACCGCTGATTTGC-3'
P30-PB2-III-F	5'-GGGGTATGAGGAGTTCACAATGGTAGG-3'
P30-PB2-III-R	5'-CCCACTGTATTGGCTTCTAATGGCC-3'
P30-HA-I-F	5'-GGCAACAGGAATGCGAAATGTACC-3'
P30-HA-I-R	5'-GCCCATATCCTCAGCATTTTCCC-3'
P30-HA-II-F	5'-CGGAGCTTCTTGTTGCCCTGG-3'
P30-HA-II-R	5'-GCAAATGTTGCACCTAATGTTGCCC-3'
P30-NA-I-F	5'-AGCAAAAGCAGGAGTAAAGATGAATCC-3'
P30-NA-I-R	5'-CCCCCGTTATACAAACATGCAGCC-3'
P30-NA-II-F	5'-GGTGGGGACATCTGGGTGACAAG-3'
P30-NA-II-R	5'-CCAAACAATGGCTACTGCTGGAGCTG-3'
P30-NA-III-F	5'-CATTGTCAGGAAGTGCTCAGCATGTCG-3'
P30-NA-III-R	5'-GAGGTCCGCCCCATCAGG-3'
P30-NP-I-F	5'-GCAAAAGCAGGGTCAATAATCACTCAC-3'
P30-NP-I-R	5'-CCATTGTCCCGATTCCTTTGACTGC-3'

Table 6 (Continued)

Primer	Sequence (5’ → 3’)
P30-NP-II-F	5’-GCCAAGCCAACAATGGTGAGG-3’
P30-NP-II-R	5’-CTGACTCTTGTGTGCTGGATTCTCG-3’
P30-M-I-F	5’-GGATGGGGGCTGTAACCAC-3’
P30-M-I-R	5’-CGGCAACAACAAGCGGGTCAC-3’
P30-M-II-F	5’-CAAATGGCTGGATCAAGTGAGCAGG-3’
P30-M-II-R	5’-GACTGTCGTCAGCATCCACAGCATTC-3’