Influenza A Virus Transmission Bottlenecks Are Defined by Infection Route and Recipient Host

Andrew Varble,1,2 Randy A. Albrecht,1,2 Simone Backes,1,2 Marshall Crumiller,3 Nicole M. Bouvier,1,4 David Sachs,5 Adolfo García-Sastre,1,2,4 and Benjamin R. tenOever1,2,*

1Department of Microbiology
2Global Health and Emerging Pathogens Institute
Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
3The Laboratory of Biophysics, The Rockefeller University, New York, NY 10065, USA
4Department of Medicine, Division of Infectious Diseases
5Department of Genetics and Genomic Sciences
Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
*Correspondence: benjamin.tenoever@mssm.edu
http://dx.doi.org/10.1016/j.chom.2014.09.020

SUMMARY

Despite its global relevance, our understanding of how influenza A virus transmission impacts the overall population dynamics of this RNA virus remains incomplete. To define this dynamic, we inserted neutral barcodes into the influenza A virus genome to generate a population of viruses that can be individually tracked during transmission events. We find that physiological bottlenecks differ dramatically based on the infection route and level of adaptation required for efficient replication. Strong genetic pressures are responsible for bottlenecks during adaptation across different host species, whereas transmission between susceptible hosts results in bottlenecks that are not genetically driven and occur at the level of the recipient. Additionally, the infection route significantly influences the bottleneck stringency, with aerosol transmission imposing greater selection than direct contact. These transmission constraints have implications in understanding the global migration of virus populations and provide a clearer perspective on the emergence of pandemic strains.

INTRODUCTION

Influenza A virus (IAV) is an important pathogen that can cause significant mortality and rapidly disseminate throughout the human population. Pandemic IAV strains can cause devastating effects on immunologically naive individuals, perhaps best exemplified by the millions of deaths caused by the 1918 H1N1 pandemic after it infected ~30% of the population (Frost, 1920; Johnson and Mueller, 2002). The emergence of a pandemic IAV is often preceded by a reassortment event, followed by rapid dissemination in the human population, as recently illustrated by swine-origin H1N1 (Smith et al., 2009). Furthermore, with the continued risk of avian-to-human H5N1 infections, and the emergence of H7N9 (Gao et al., 2013), there is significant impetus in gaining a greater understanding into the dynamics of IAV transmission.

IAV has been shown to transmit through direct contact, large droplets, and aerosols (Brankston et al., 2007; Tellier, 2009). The relative contribution of each mode of transmission remains contested, but it is clear that the aerosol route is a successful means of virus dissemination among the human population (Cowling et al., 2013). Given this, attempts to quantify aerosol secretions from infected patients have yielded data suggesting as many as $10^5$ viral copies can be excreted over a 30 min period (Milton et al., 2013). In addition, aerosol administration to volunteers found the minimal infectious dose to be very low, estimating that fewer than ten virions could account for infection (Alford et al., 1966). These numbers have also been found to be comparable in the ferret IAV model (Gustin et al., 2011). While these data illustrate some important characteristics of virus transmission and give insight into how the virus spreads efficiently on a global scale, they fail to address the physiological dynamics that occur in nature. This information is critical to better understand pandemic emergence and global virus dissemination.

Defining the characteristics of the viral population during transmission has important implications in the control of disease and the evolutionary path of the virus. The high error rate of RNA virus’ RNA-dependent RNA polymerases leads to a naturally maintained level of genetic variation within the population, sometimes referred to as a viral quasispecies. This diversity is necessary for the fitness of viral populations, as it allows the virus to quickly move into new genetic space following different selective pressures (Vignuzzi et al., 2006). Repeated artificial bottlenecks in viral populations have been demonstrated to severely restrict viral fitness (Duarte et al., 1992), and natural bottlenecks have been observed during a wide variety of virus transmission and dissemination events, including human immunodeficiency virus, Venezuelan equine encephalitis virus, polio virus, hepatitis C virus, and various plant viruses (Derdeyn et al., 2004; Forrester et al., 2012; Wang et al., 2010). Furthermore, studies have utilized genetic tags to track polio and VEEV populations during these events (Forrester et al., 2012; Lauring and Andino, 2011; Pfeiffer and Kirkegaard, 2006). Here we present an approach
RESULTS

Design of Barcoded IAV Library

IAV is a member of the family Orthomyxoviridae, and segments seven and eight of its genome undergo splicing as a means to encode multiple proteins from a single segment (Shaw and Palese, 2013). We exploited a modified segment eight that introduces a noncoding intergenic region between the NS1 and NS2 (also called the nuclear export protein, or NEP) sequences (Varble et al., 2010) and inserted a 22 nucleotide (nt) barcode into this site (Figure 1A). To determine if modification and insertion of barcodes into the viral genome were genetically neutral, we rescued this virus in the H1N1 2009 pandemic background (A/California/04/2009) and compared its growth to a virus containing the wild-type (WT) NS segment. Multicycle replication in human lung epithelial cells confirmed equivalent growth of these viruses and validated this approach to study IAV populations in real time and under physiological conditions (Figure 1B).

To monitor bottlenecks during IAV transmission, we individually rescued over 100 viruses with unique 22 nt, GC-content-matched barcode identifiers in the intergenic region of the NS segment. Multicycle growth curve analysis of a subset of these viruses demonstrated tightly grouped replication levels with only one significant outlier which grew poorly, possibly as a result of impacting splicing (Chua et al., 2013) (see Figure S1A available online). Given the overall lack of a replication-based phenotype upon barcode insertion, we quantified our virus population and combined equivalent levels of each unique virus. This viral population was deep sequenced to confirm the distribution and composition of the viral clones present in the library. Sequencing revealed the presence of a stoichiometric balance between each barcoded virus, with no member represented at over five percent of the population (Table S1). Furthermore, this deep sequencing strategy reproducibly generated comparable barcode profiles within a particular population across duplicate samples, suggesting this technique to be an accurate surrogate measure of quasispecies (Figures S1B–S1E).

Propagation In Ovo, but Not In Vitro, Results in Sequence-Specific IAV Bottlenecks

The barcoded virus library was propagated in both cell culture and eggs to ascertain the replication characteristics of the mixed viral population. IAV infection of Madin-Darby canine kidney (MDCK) cells, the standard culturing system for growing the virus in vitro, demonstrated no bottleneck during virus amplification following a low multiplicity of infection (moi). This was evident by a uniform and reproducible distribution of amplified viral clones across three individual sets of experiments (Figure 2A; Table S1). This phenotype was not restricted to MDCKs, as these trends were also observed during propagation in a human epithelial cell line (Figure 2B; Table S1). Conversely, infecting embryonated-chicken eggs with an equivalent dose of the mammalian IAV-based barcoded library encountered stringent bottlenecks, with only 5–13 viral clones being successfully amplified in each egg (Figure 2C; Table S1). Furthermore, this bottleneck occurred in a seemingly nonspecific manner, as a diverse cohort of clones emerged in different sets. To confirm that the amplification of these specific viruses was not due to inherent restraints or replicative advantages based on barcode, both amplified and nonamplified viruses were individually compared in vitro and in ovo. These infections demonstrated no aberrations that would explain the selection in ovo (Figures S1A and S2A). To explore the possibility that selection was taking place postinoculation, the hemagglutinin (HA) segments of individual viruses were analyzed. All viruses sequenced obtained amino acid changes in residues previously implicated taking place postinoculation, the hemagglutinin (HA) segments of individual viruses were analyzed. All viruses sequenced obtained amino acid changes in residues previously implicated in sequence-specific IAV bottlenecks.
the virus to acquire mutations that allow for optimal viral entry in the egg. Although we cannot rule out the possibility of beneficial mutations in other segments, the changes detected in the HA suggest a scenario whereby the individual viruses that first acquire increased avian receptor specificity successfully outcompete the remaining viruses present in the inoculum.

Transmission Results in Sequence-Independent IAV Bottlenecks

As the barcoded viruses effectively identified conditions in which bottlenecks occurred, we sought to examine transmission dynamics in vivo. First, we examined transmission in guinea pigs, as this model recapitulates many of the characteristics of human virus spread (Lowen et al., 2006). To this end, virus-donor guinea pigs were intranasally infected with the IAV barcoded library, and then, 24 hr later, each donor animal was paired with a naive virus-recipient guinea pig. Donor and recipient guinea pigs were housed separately in neighboring cages so that contact between animals was minimized and thus virus transmission presumably occurred by droplet or aerosol routes. Animals were monitored by nasal wash in which transmission was determined by standard plaque assay and viral populations were defined by next-generation sequencing. Importantly, characterization of the IAV barcodes in the inoculated animals demonstrated a profile comparable to that observed in vitro, with approximately three-quarters of the viral library readily detectable from the nasal wash (Table S2).

Upon evaluation of the recipient infections (RIs), we found three out of the four guinea pigs developed positive titers for IAV by 4 days postexposure (Figure 3A). Remarkably, the three infected contact animals all demonstrated evidence for a stringent bottleneck, with only two to five clones detected at day 6 (Figure 3B; Table S2). Overall, these data clearly illustrate a significant bottleneck during IAV transmission and a resetting of the viral population.

To determine if the bottleneck was (1) independent of virus genetics and (2) occurring at the level of donor or recipient, we inoculated a single animal and caged three naive animals with it to monitor multiple transmission events to different recipients from a single donor. All three recipient animals developed positive nasal wash viral titers within 4 days of caging (Figures 4A and 4B). Strikingly, the barcode profile between the three recipient animals was markedly different despite identical exposures (Figures 4C and 4D; Table S3). These data suggest that the observed bottlenecks among susceptible hosts can be independent of virus genetics and occur at the level of the recipient, although the possibility that some selection also takes place at virus secretion still exists. Furthermore, we observe increases of barcode diversity at an average of 25 barcodes per animal over time in the recipient animals, suggesting ongoing transmission events or amplification of virus populations from the initial contact (Table S3).

Route of Transmission Affects Bottleneck Stringency

To more formally parse out the contribution of contact versus airborne transmission bottlenecks, we repeated these studies with three groups of ferrets. Three donor ferrets were directly inoculated with the IAV barcode library (directly inoculated, DI).
denoted as DI, d6 RI, or d8 RI, respectively. See also Table S2.

After 24 hr, each DI ferret was cocaged with one recipient ferret (contact infection, CI) and physically separated from another recipient ferret (airborne infection, AI) (Figure 5). Intranasally inoculated DI ferrets shed virus at levels sufficient to infect all recipient ferrets (airborne infection, AI) (Figure 5). Intranasal inoculation or nebulized virus. Similar to the results obtained during ferret and guinea pig transmission experiments, we first sought to determine if, as we observed in eggs, selective pressure on the HA segment was responsible. We sequenced viruses from direct-, contact-, and airborne-infected animals and observed disparate mutations as opposed to the complete viruses from direct-, contact-, and airborne-infected animals and observed disparate mutations as opposed to the complete replication of specific residues that were observed in eggs.

When observing transmission in both guinea pigs and ferrets, infection through the airborne route results in a significantly reduced number of clones detected in the recipient animals (Figure S3B). To further explore this observation, we infected mice with equivalent doses of the virus library through either intranasal inoculation or nebulized virus. Similar to the results obtained during ferret and guinea pig transmission experiments, inoculation through the airborne route results in more significant bottlenecks and reduction in the number of barcodes detected (Figures S3C–S3E; Table S4). These data suggest that airborne transmission imposes one of the greatest bottlenecks encountered by the virus population even when secreted at equal opportunity for infection.

The Probability and Stochastic Nature of Transmission

If transmission were considered to be stochastic, it would imply that each virus has a nonzero probability of transmission. Conversely, it is possible that some viruses have developed mutations that increase fitness and confer a probabilistic advantage of transmission relative to their peers. If this were the case, the distribution of successful transmissions would differ statistically from a population of viruses in which all had equal probability of transmission. In the scenario where each virus has equal opportunity to infect the recipient animal, the resulting number of successful transmissions for the population of viruses would form a binomial distribution. However, not all viruses were present in all sources, and the proportions of each barcode were not perfectly balanced. Therefore, the number of potential transmission events, and probability of transmission, differed between viruses.

To investigate the stochastic nature of the bottleneck, we first noted that the probability of a successful transmission of a virus in a given population is related to the initial proportion in which the virus is present. We established a relationship between the initial virus proportion in the donor animal and the proportion of successful transmission events in all guinea pig and ferret experiments (Figure 6A). We then modeled this correlation with a step-response function depicted by the dashed red curve, bounded between transmission rates of zero and unity over the range of possible initial proportions, and representative of the data. This relationship was used to determine the likelihood of observing the collected data, provided the assumption that the only contributing factor to virus transmission is the initial proportion of the virus present. Using initial conditions identical to those found in the laboratory (number of viruses, number of possible transmission events), a Monte-Carlo simulation was used to generate a large set (10^6) of transmission outcome distributions.
in which each virus transmitted with a probability as determined by the model. We then observed the likelihood of observing the laboratory distribution in the context of our simulated distributions. Any one distribution in the set can be compared to the mean (expected) distribution via a distance measure $d$ defined as the sum-of-squares difference between distributions (Figure 6B). The distance between the laboratory distribution and the average simulated distribution fell into the 30th percentile (Figure 6C), indicating that the collected data is representative of a purely stochastic bottleneck in which the only contributing factor is the initial virus proportion relative to the population.

Some overlap among barcodes and independent transmission events was observed in both ferrets and guinea pigs (Table S2, Table S3, and Table S4). To ensure this was not due to an inherent growth advantage as a result of the barcode sequence, these shared clones were compared to randomly selected viruses that did not transmit (Figure S1). These data clearly demonstrated that the barcodes that were detected in multiple samples did not convey a replication advantage (Figure S1). It therefore appears the prevalence of these specific clones was due to their modest overrepresentation in the original library, which is also supported by a positive correlation between virus proportion and likelihood of transmission (Figure 6A). Together, these data suggest that increased titers in the host contribute to the propensity for transmission.

**Virus Proportions in Nasal Wash Correlate with Transmission**

Previous studies have proposed that the site of replication within the respiratory tract determines the transmissibility and pathogenicity of IAV (van Riel et al., 2010). Specifically, the inability to bind α-2,6 linked sialic acid, and therefore replicate in the upper respiratory tract of mammals, is thought to be one of the main constraints for human-to-human transmission of avian IAV strains (Shinya et al., 2006; Tumpey et al., 2007; van Riel et al., 2006). To determine whether our system accurately reflected the proposed model in which viruses in the upper respiratory tract are responsible for transmission, we profiled virus libraries in both the nasal wash (upper respiratory tract) and bronchus tissues (lower respiratory tract) and compared these populations to those transmitted by both direct contact and airborne transmission in ferrets.

During direct CI, we only observe a significant correlation between replication proportions in the nasal washes and the

---

**Figure 4. Transmission from Single Donor to Multiple Recipient Guinea Pigs**

(A) Viral titers in the nasal washes of guinea pigs collected at indicated days postinfection (dpi) are reported as plaque-forming units (pfu/mL). Solid line depicts inoculated animals (DI), whereas dashed lines denote titers from three naive recipient guinea pigs placed into direct contact with the donor 1 day postinfection (CI). A–D indicate individual caged animals.

(B) Plots representing viral barcodes present in viral populations. Each color depicts a unique barcode whose relative proportion corresponds to its abundance in the virus population in the indicated sample. “Library” denotes starting virus material, whereas direct infection on day 2 or recipient infections on days 4 or 7 are denoted as DI, d4 DI, or d7 DI, respectively. See also Table S3.
likelihood of virus transmission, suggesting that viruses replicating in the bronchus do not readily transmit (Figures 7A and 7B; Table S4). As it is not possible to distinguish the source of a given transmission event when the corresponding barcode is detected in both upper and lower respiratory tracts, it should be noted that there are no examples of a bronchus-only population that successfully transmits in contrast to eight nasal wash-only transmission events (Table S4). These trends are also observed during airborne transmission (Figures 7C and 7D; Table S4), suggesting that replication in the upper respiratory tract is the source of virus for both direct and airborne transmission events.

Furthermore, we reran the Monte-Carlo simulation using this subset of samples, with the intent of comparing the relative abilities of measurements from both areas to predict virus progression. In both cases, we again used the measured relative proportion of each virus from the source animal as a predictor for transmission rate, and estimated the statistical likelihood of observing the final proportions found in the target. The bronchus provided transmission results that were highly unlikely to be observed, assuming the model parameters estimated from the virus proportions in the bronchus (Figure S4A; a = 6.1, p = 0.003). Conversely, samples from the nasal washes provided transmission rates consistent with those generated by the model (Figure S4B; a = 12.7, p = 0.73).

**DISCUSSION**

Here we apply genetic barcoding of the IAV genome as a tool to define bottlenecks encountered during virus dissemination in vitro, in ovo, and in vivo. Bottlenecks encountered in ovo, during a period of adaption, were found to be the result of virus receptor remodeling for the avian host. This type of selective pressure on an adapting avian HA within a mammalian host has also recently been reported (Wilker et al., 2013). Conversely, while we find that adapted IAV strains also undergo strict bottleneck events during mammalian transmission, these occur independently of HA genetics acting on particular virus
populations. These results provide a deeper understanding into the dynamics of IAV quasispecies during transmission events by suggesting that as few as two barcodes can serve as the founder viruses to initiate productive infections. Furthermore, we find transmitted viruses appear to originate from the upper respiratory tract and hypothesize that these virions are expelled at levels related to their overall proportion at this site of infection. Lastly, we hypothesize that viruses from the donor are excreted with equal opportunity for transmission and that the observed bottlenecks occur at the level of the recipient.

How IAV retains its relative fitness despite these repeated bottlenecks is unclear. It is possible viral fitness is, in fact, maintained because of a strong purifying selection of transmitted viruses, resulting in productive infection of only the most fit viruses, or, alternatively, that some individuals (or conditions) do not impose the same restrictions on incoming virus and subsequently serve as local virus reservoirs and superspreaders. Additionally, these findings may help explain the relatively low virus reproductive value ($R_0$), between 1 and 2, that has been described for IAV (Dorigatti et al., 2013; Fraser et al., 2009; White et al., 2009). Furthermore, it has recently been noted that some circulating H5N1 strains are only three amino acid substitutions away from achieving airborne transmission in ferrets (Herfst et al., 2012; Imai et al., 2012; Russell et al., 2012). These mutations were predicted to have the ability to arise concurrently in an infected mammalian host, but only at proportions ranging from $10^{-7}$ to $10^{-11}$. Our findings indicate that dissemination of viruses at these levels is unlikely and offers an explanation as to why, despite significant numbers of human infections, a mammalian transmissible H5N1 has not yet emerged.

Lastly, our data illustrate that in order for a virus like H5N1 to develop into a pandemic it must pass through two distinct bottlenecks. As demonstrated previously, the virus must first acquire a specific set of mutations and/or reassortments to adapt and allow for optimal replication in the new host (Russell et al., 2012). Second, the virus must also overcome the sequence-independent bottlenecks we have defined that occur during transmission between susceptible hosts. Infections following direct contact with the host are more efficient and maintain a greater proportion of the original virus population when compared to airborne transmission events. Given these results, preventative measures such as avoiding direct contact with infected individuals may help curb the emergence of future pandemics.

**EXPERIMENTAL PROCEDURES**

**Virus Design and Rescue**

The NS segment was split as previously described (Varble et al., 2010). The 22 nt barcoded sequence was amplified from the 3' arm of a shRNA library (Silva et al., 2003), along with 100 base pairs of common flanking sequence. This sequence was inserted into a BstEII site positioned in the intergenic region between NS1 and NS2, and standard reverse genetics were used to individually rescue barcoded viruses. Viruses were propagated on MDCK cells, then quantified using hemagglutination assays and combined equal levels to construct the viral library. Egg-adapted viral library was constructed by injecting 10,000 plaque-forming units of original virus library into 20 10-day-old specific pathogen-free embryonated chicken eggs (Charles River), and allantoic fluid was harvested 48 hr postinfection. Allantoic fluid from all eggs was then combined at equal volume to create egg-adapted virus library.

**Tissue Culture and Virus Infections**

A549 and MDCK cells were cultured in DMEM media supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. For virus infections, indicated cell lines were incubated with virus at the indicated moi with PBS supplemented with 0.3% BSA (MP Biomedicals), 10 mM CaMg, and penicillin/streptomycin for 1 hr, washed with PBS, and serum-free DMEM media supplemented with 0.3% BSA and 0.2 μg/ml (A549s) or 1 μg/ml (MDCKs) of TPKC trypsin was added. Multicycle growth curves were performed in biological triplicates with 100 μl of supernatant removed at indicated time points and titers assessed by plaque assay. Ten-day-old specific pathogen-free
embryonated chicken eggs (Charles River) were infected at the indicated doses, and allantoic fluid was harvested 48 hr postinfection.

Deep Sequencing and Data Analysis of Virus Libraries
To monitor viral populations, Superscript III One-step RT-PCR (Invitrogen) was used with specific primers to the NS segment. Nested PCR was then used with barcoded Illumina linkers to amplify the barcoded region. Deep sequencing samples were analyzed on the Illumina MiSeq sequencing platform. A total of 15–30 samples were multiplexed per run with an average of approximately 225,000 barcode reads per sample. Barcoded reads were extracted initially by searching for a 19 nt sequence that precedes all 22 nt barcodes. The following 22 nt sequence for each of these reads was then matched against each barcode and aggregated. Limit of detection for viral populations was set at 100 reads, as this represented the upper end of virus background detected in mock samples. Propagation and transmission experiments were visualized using Matlab.

Animal Transmission Experiments
Male Fitch ferrets (gibs) at 5 months of age were purchased from Triple F Farms (Sayre, PA) and confirmed to be serologically naive for currently circulating H1N1 and H3N2 IAV strains and influenza B virus strains. Influenza virus infection and transmission studies with ferrets were conducted as described elsewhere (Baker et al., 2013; Seibert et al., 2010). Ferrets were anesthetized by intramuscular injection of ketamine (10 mg/kg) and xylazine (2 mg/kg). For transmission studies, a seronegative virus-donor ferret was directly infected by the intranasal route with 10,000 plaque forming units (pfu). At 24 hr postinoculation naive contact and airborne recipient ferrets were cohoused with the donor ferret. Nasal washes were collected from anesthetized ferrets at 2, 4, and 6 days postinfluenza infection. Female Hartley strain guinea pigs at 4–5 weeks of age (300–350 g) were obtained from Charles River Laboratories (Kingston, NY). Transmission studies of influenza viruses among guinea pigs were conducted as described previously (Chou et al., 2011; Lowen et al., 2008). Briefly, for all transmission studies, guinea pigs were anesthetized by intramuscular injection of ketamine (30 mg/kg) and xylazine (5 mg/kg). The inoculated virus-donor guinea pigs were directly infected by the intranasal route with 10,000 pfu, and nasal washes were collected from all anesthetized guinea pigs at 2, 4, 6, and 8 days postinoculation. In contact transmission experiments, three naïve recipient guinea pigs were cohoused with the directly infected donor guinea pig, in a single cage, at 24 hpi. In respiratory droplet transmission experiments, a directly infected donor guinea pig and a naïve recipient guinea pig were placed in separate cages in which one side was replaced with a wire mesh panel. Donor and recipient cages were placed side by side, with wire mesh panels opposed, so that air could flow freely between cages but direct contact between animals was precluded, allowing transmission to occur only by droplet spray or aerosol routes. To compare intranasal infection to AI, a virus inoculum containing the barcoded virus library was prepared. Balb/C mice were anesthetized with ketamine/xylazine prior to either exposure to virus nebulized by an inExpose Bench-top inhalation
exposure apparatus (SCIREQ Scientific Respiratory Equipment Inc.) or intra-nasal infection. Anesthetized mice were exposed to an equivalent of 10,000 pfu of nebulized virus following manufacturer’s instructions and suggested equations for calculating delivered virus dose or infected with 10,000 pfu intra-nasally. Lungs were harvested 3 days postinfection. All ferret, mice, and guinea pig studies were reviewed and approved by the institutional animal care and use committee.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and four tables and can be found with this article at http://dx.doi.org/10.1016/j.chom.2014.09.020.

ACKNOWLEDGMENTS

We wish to thank Eric Jaffe for portions of the guinea pig work. This research was partially supported by Center for Research on Influenza Pathogenesis (CRIP), an NIAID-funded Center of Excellence for Influenza Research and Surveillance (CEIRS), contract number HHSN266200700010C (to R.A.A., N.M.B., A.G.-S., and B.R.t.), and an R01 awarded to B.R.t. (grant number A1093571).

REFERENCES


