Viral expansion after transfer is a primary driver of influenza A virus transmission bottlenecks

Katie E. Holmes[1], David VanInsberghe[1], Lucas M. Ferreri[1], Baptiste Elie[2], Ketaki Ganti[1], Chung-Young Lee[3], Anice C. Lowen[1,4*]

[1]Department of Microbiology and Immunology, Emory University School of Medicine, Atlanta, GA, USA
[2]MIVEGEC, CNRS, IRD, Université de Montpellier, Montpellier, France
[3]Department of Microbiology, School of Medicine, Kyungpook National University, Jung-gu, Daegu, Republic of Korea
[4]Emory Center of Excellence for Influenza Research and Response (CEIRR), Atlanta, GA, USA

*anice.lowen@emory.edu

ABSTRACT

For many viruses, narrow bottlenecks acting during transmission sharply reduce genetic diversity in a recipient host relative to the donor. Since genetic diversity represents adaptive potential, such losses of diversity are though to limit the opportunity for viral populations to undergo antigenic change and other adaptive processes. Thus, a detailed picture of evolutionary dynamics during transmission is critical to understanding the forces driving viral evolution at an epidemiologic scale. To advance this understanding, we used a novel barcoded virus library and a guinea pig model of transmission to decipher where in the transmission process diversity is lost for influenza A viruses. In inoculated guinea pigs, we show that a high level of viral genetic diversity is maintained across time. Continuity in the barcodes detected furthermore indicates that stochastic effects are not pronounced within inoculated hosts. Importantly, in both aerosol-exposed and direct contact-exposed animals, we observed many barcodes at the earliest time point(s) positive for infectious virus, indicating robust transfer of diversity through the environment. This high viral diversity is short-lived, however, with a sharp decline seen 1-2 days after initiation of infection. Although major losses of diversity at transmission are well described for influenza A virus, our data indicate that events that occur following viral transfer and during the earliest stages of natural infection have a predominant role in this process. This finding suggests that immune selection may have greater opportunity to operate during influenza A transmission than previously recognized.

INTRODUCTION

The high mutation rate characteristic of RNA viruses [1, 2], coupled with genetic recombination [3, 4] and/or reassortment [5] of segmented genomes, enables constant production of viral variants [6-8]. In turn, this variation provides the substrate for viral evolution [9-11]. For some viruses, including influenza A virus and SARS-CoV-2, viral evolution and viral spread through host populations occur on a similar timescale, such that each potently shapes the other [12]. Understanding viral evolution is therefore crucial in our efforts to control outbreaks.

Two major evolutionary processes mediate viral change: selection and genetic drift [13]. Selection is a deterministic evolutionary force by which individuals with advantageous traits are more likely to survive and reproduce. In turn, alleles from those individuals will increase in frequency over multiple generations. Conversely, genetic drift is a stochastic process whereby changes in allele frequency occur due to random events. Selection and genetic drift both function to remove genetic diversity from a population: one through the increase of alleles under selection and the other through stochastic loss of alleles from a population. Their relative influence in any evolutionary system varies. In general, when a large number of individuals are able to reproduce, selection acts more efficiently because the probability of any given variant becoming extinct is low. However, when very few individuals are able to reproduce, the probability of alleles being stochastically lost from the population is higher, making selection less efficient. In this way, selection and genetic drift are opposing forces: where one is strong, the other is weak. While they both work to lessen genetic diversity, they have fundamentally different impacts on the population's evolutionary trajectory. For influenza viruses, the interplay between selection and genetic drift is not well understood.

A reduction in viral population size during transmission has been documented for many viruses [14-18]. Termed a transmission bottleneck, this effect is characterized by the establishment of a new infection by only a small number of genomes derived from a large source population. Transmission bottlenecks typically impose a stochastic loss of diversity, rather than a selective one in which successful variants are those with the highest fitness. When a stochastic bottleneck is active, even highly adaptive alleles may go extinct.

As for other viruses, loss of diversity during influenza A virus transmission has been documented. In humans with naturally occurring infections, as few as one to two viral genomes from an infected individual established in an uninfected individual [19]. Studies using influenza A virus populations with inserted neutral barcodes similarly revealed narrow transmission bottlenecks in animal models [20, 21]. This work also demonstrated that aerosol transmission caused a more stringent loss of diversity than transmission where all modes were possible [20, 21]. More relaxed bottlenecks have been reported in natural hosts aside from humans, suggesting roles for host biology and/or modes of transmission in shaping species-specific between-host dynamics [22, 23]. Finally, avian influenza A virus transmission in mammals has been shown to involve a tight bottleneck with reductions in viral diversity attributed to either stochastic or selective forces [24-26].

While it is known that diversity is lost during influenza A virus transmission, it is unclear at which point – in the donor host, the environment, or the recipient host – this takes place. The specific stage at which diversity is lost is likely to define the potential for selection to act during transmission. We therefore made an influenza A virus library with 4096 potential barcodes to monitor the fate of many unique viral lineages within experimentally infected guinea pigs and through the course of transmission to contacts. Our data reveal that viral diversity remains high in inoculated animals throughout the course of infection, and many viral genomes are transferred to naïve animals exposed by aerosols or direct contact. However, a severe

87 bottleneck occurs in contact animals 1-2 days after the initiation of infection, such that few
88 lineages are sustained in the recipient. Thus, in our system, losses of diversity are primarily
89 driven by events that occur during the expansion of infection, not during the process of donor-
90 to-recipient transfer itself. Importantly, this phase of expansion could be a previously
91 unrecognized opportunity for immune selection to operate.
92

93 RESULTS

94 **Generation of a Barcoded Influenza A Virus Library**

95 We designed a genetic barcode for influenza A/Panama/2007/99 (H3N2) virus (Pan/99) with
96 the goals of avoiding attenuation of viral replication and producing a highly diverse viral
97 population with minimal fitness differences among variants. To achieve this, we introduced
98 polymorphisms within the native neuraminidase (NA) sequence of Pan/99 rather than inserting
99 a foreign sequence. Twelve barcode sites within a 50-nucleotide region of the NA segment were
100 selected based on synonymous single nucleotide variants found in H3N2 subtype influenza A
101 viruses circulating in humans from 1994 to 2004. At each of these 12 sites, two nucleotides
102 were possible, resulting in 4096 unique barcodes (**Figure 1A**). We named the gene segment
103 modified in this way Pan/99 NA-BC.
104

105 To evaluate barcode diversity in the reverse genetics plasmid and the passage 1 (P1) viral stock
106 carrying Pan/99 NA-BC, we subjected the barcode region to next-generation sequencing. No
107 barcode was found to be dominant in the plasmid library or the virus stock, and no nucleotide
108 was dominant at any site within the barcode (**Figure 1B, C**). An aberrant nucleotide was
109 detected at the twelfth barcode site in the plasmid preparation but was not carried through to
110 the virus stock. To quantify diversity, we applied the Shannon Diversity Index (*H*), which
111 considers the richness (i.e., the number of species present) and the evenness (i.e., the
112 abundance of the present species) in a community [27]. To calculate *H*, each unique barcode
113 was taken to represent a species. The measured diversity in the plasmid and virus stocks was
114 8.07 and 7.90, respectively (**Figure 1D**), near to the theoretical maximum (H_{max}) for this system
115 of 8.32 and revealing little loss of diversity during viral recovery from cDNA. To determine
116 whether the barcode in the NA segment altered the growth phenotype of the Pan/99 virus,
117 multi-cycle replication was evaluated in MDCK cells. No significant differences between wild
118 type and Pan/99 NA-BC viruses were detected (**Figure 1E**).
119

120 **Modeling Transmission of Pan/99 NA-BC Virus in Guinea Pigs**

121 To evaluate viral dynamics within and between hosts, we inoculated guinea pigs with 5×10^4
122 plaque-forming units (PFU) of Pan/99 NA-BC virus intranasally [28]. After 24 hours, we placed a
123 naïve animal with each inoculated animal in either direct or aerosol contact (**Supplemental**
124 **Figure 1A**). Three independent experiments were performed, each including four direct contact
125 and four aerosol contact pairs. Viral shedding, assessed by plaque assay of daily nasal lavage
126 samples, was similar for all inoculated animals (**Supplemental Figure 1B**). Transmission
127 efficiency varied from 75% to 100% for pairs in direct contact and from 50% to 100% for pairs in
128 aerosol contact. The daily nasal lavages collected from both inoculated and exposed animals

129 furnished a valuable set of samples with which to investigate viral population dynamics and the
130 drivers of the influenza A virus transmission bottleneck.
131
132 **Stochastic Effects are Not Pronounced in Inoculated Animals**
133 To determine how viral diversity changed during infection in inoculated animals, we evaluated
134 the barcodes present in nasal lavage samples by next generation sequencing. Many barcodes
135 were detected throughout the course of infection in all directly inoculated animals (**Figure 2**
136 and **Supplemental Figure 2**). Using Shannon's index (*H*), we found that viral population diversity
137 declines gradually over time in most inoculated animals (**Figure 3** and **Supplemental Figure 3**).
138 Changes in diversity are furthermore mirrored by changes in both barcode richness and
139 evenness, indicating that declines in diversity are driven both by loss of barcode species from
140 the within-host population and skewing of the relative frequencies of individual barcodes that
141 comprise the population (**Figure 3** and **Supplemental Figure 3**).
142
143 To evaluate how populations detected on a given day relate to those present at prior and
144 subsequent time points in the same animal, we applied two measures of dissimilarity: Bray-
145 Curtis and Jaccard. Both statistics gauge the compositional disparateness between two samples
146 but differ in that Jaccard considers only the extent to which two populations differ in the
147 species present, while Bray-Curtis considers both differences in the species present (or not
148 present) and differences in the abundance of each species. These analyses revealed relatively
149 low dissimilarity across longitudinal samples collected from a given inoculated animal
150 (**Supplemental Figure 4**). The barcode composition of each sample was typically most closely
151 related to those of the immediately preceding and following days.
152
153 To aid in relating these results to those obtained from natural infections, which are
154 characterized by markedly lower viral genetic diversity than examined here, we binned the
155 4096 potential barcodes into four classes based on the identity of the first two nucleotides in
156 the barcode. We then evaluated the dynamics of these four sub-populations over time (**Figure 4**
157 and **Supplemental Figure 5**). In inoculated animals, the frequency of the four sub-populations
158 was found to be highly stable across sampling days. The gradual nature of observed shifts in
159 barcode composition over time indicates that infection in animals inoculated with a large viral
160 population is shaped only minimally by stochastic effects.
161
162 **High Initial Diversity in Exposed Animals Precedes a Sharp Decline**
163 To determine how viral diversity changed between donor and recipient animals, barcodes
164 present in nasal lavage samples of recipients were analyzed. Strikingly, at the earliest
165 timepoint(s) positive for infectious virus in recipients, many barcodes are present, signifying
166 transfer of an appreciably large viral population through the environment (**Figure 2** and
167 **Supplemental Figure 2**). This observation is true of both aerosol-exposed and direct contact
168 animals. Populations present early in exposed animals are characterized by high diversity,
169 richness, and evenness (**Figure 3** and **Supplemental Figure 3**). Based on a Jaccard analysis,
170 these populations show relatedness in the species present to viral populations replicating
171 within the corresponding donor animal (**Supplemental Figure 4**).
172

However, the observed high viral diversity in exposed animals is short-lived, with a sharp decline seen within 1-2 days of the initiation of infection (**Figure 2** and **Supplemental Figure 2**). Both richness and evenness contribute to the precipitous drop in diversity seen within exposed animals (**Figure 3** and **Supplemental Figure 3**). Accordingly, the Bray-Curtis index, which considers both shared species and their comparative abundances, reveals high dissimilarity between populations present in recipient animals on the first positive day and at subsequent time points compared with the same time points in donor animals (**Supplemental Figure 6**). Of note, barcode composition in exposed animals remains highly consistent after the initial 1-2 days of viral positivity (**Figure 2** and **Supplemental Figure 2**), indicating that the few viral lineages that penetrate the bottleneck persist through the remainder of the acute infection. The barcodes that do persist furthermore differ across independent exposed animals (**Figure 2** and **Supplemental Figure 2**), excluding the possibility that a subset of barcodes carry a fitness advantage.

To assess the potential for these dynamics to be driven by selection acting on *de novo* mutations outside the barcoded region, we performed whole-genome sequencing of a subset of samples. Specifically, viral genomes collected on the first or second day of positivity in exposed animals from experimental replicates 1 and 2 were sequenced in full. Few *de novo* mutations were detected; none were found on successive days in the same animal; and all non-synonymous changes remained below a 20% frequency (**Supplemental Figure 7**). Thus, reductions in viral diversity observed in exposed animals were not driven by selective sweeps of *de novo* variants. In addition, a focused analysis of whole genome sequencing reads that spanned the barcode region was used to validate results obtained from amplicon sequencing. Although the small number of reads meeting the criterion to span the entirely of the barcode limited this analysis, high-level observations could be confirmed: first positive time points showed multiple barcodes present and subsequent time points revealed the same predominant barcodes in both sequencing datasets.

Taken together, these findings signify that many viral genotypes replicating in inoculated animals are transferred to exposed animals. Some stochasticity in this transfer is apparent in that genotype abundance in the recipient is not tightly linked to that in the donor. However, strong stochastic effects are seen within the recipient. Very few viral genotypes ultimately establish infection, signifying that processes occurring during early expansion of the viral population limit diversity in newly infected hosts.

DISCUSSION

We used a barcoded influenza A virus system to generate a high-resolution view of changes in the composition and diversity of viral populations as they expand, contract, transmit onward, and establish infection in a new host. Our data suggest that viral evolutionary dynamics through a transmission event comprise two distinct stages: inter-host transfer and intra-host establishment. The bottleneck acting during the first stage can be very loose, allowing many viral genotypes to pass to the new host. Conversely, the bottleneck acting at the second stage is highly stringent such that few genotypes predominate the established within-host population.

The evolutionary implications of this two-stage process are potentially great. The existence of a loose bottleneck between hosts could enable selection to act efficiently on the transferred viral population *prior to* a stochastic contraction of diversity imposed by the process of viral establishment. Indeed our data offer a potential solution to a long-standing conundrum in evolutionary virology, that of disparate evolutionary dynamics at the individual and global population scales [29].

Influenza virus evolution at large geographic and temporal scales is characterized by a very clear pattern of positive selection: on a recurring basis, antigenically distinct variants sweep the global viral population, driving epidemic spread [30-32]. In stark contrast, the within-host evolutionary dynamics of influenza viruses show strong genetic drift and purifying selection; positive selection has rarely been observed [19, 33-35]. A similar dichotomy is true for SARS-CoV-2 at within-host and population scales [36-39]. Our data suggest that these scales may be linked by a transmission process in which selective forces have an opportunity to act before potent stochastic forces within hosts. For example, antibodies present at mucosal surfaces could act on an antigenically diverse incoming viral population, mediating antigenic selection. Only positively selected variants would then be available to pass through the stochastic bottleneck associated with population expansion. Experiments performed in pre-immune hosts are needed to formally test this model.

The reduction in diversity between virus populations in donor animals and those that become established in recipients seen here is highly consistent with prior studies carried out both in human cohorts and experimental animals [19-21, 24]. Our conclusion that early viral dynamics in the recipient make a major contribution to this loss of diversity is, however, novel. This pattern was likely not apparent in prior work simply due to the time points analyzed. In human cohort studies, sample collection is typically triggered by the onset of symptoms reported by the study subject, such that the very early stages of infection are often not sampled [19]. Similarly, previous studies of the influenza A virus transmission bottleneck in experimental models did not characterize viral populations present early after transmission [20, 24].

In contrast to recipient animals infected through transmission, we found only modest reductions in diversity in donor animals during infection. This finding may be explained by the relatively high viral load delivered during inoculation and/or differences in the sites of viral deposition given intranasal inoculation vs. natural exposure.

The observation that many viral genomes were conveyed through the environment during both direct and aerosol exposures is somewhat surprising in light of prior studies that compared different types of exposure [20, 24]. This result was also unanticipated based on the potential for expelled aerosols to be diluted in air [40]. While our data do not indicate how transmitted viruses are transferred, efficient delivery of multiple viruses to a recipient may be a result of their co-incorporation into aerosol particles. In addition, or alternatively, efficient transfer of influenza A viruses may rely on close proximity between donor and recipient or extended exposure times [41], which were both features of our experimental design.

260 The reduction in neutral diversity observed early in the course of infection in exposed animals
261 indicates that viral dynamics early after natural infection are characterized by strong stochastic
262 effects. This stochasticity could have multiple drivers. One important consideration is whether
263 the viral genotypes detected early in recipients' nasal lavage samples are replicating. Although
264 some early samples in our dataset are both high diversity and high titer, suggestive of active
265 replication of many barcodes, our sequencing approach cannot distinguish viral genomes
266 present within virions from those undergoing active replication within cells, leaving uncertainty.
267 Viruses delivered to a new host might fail to initiate replication due to mechanical trapping,
268 such as in mucus within the nasal cavity [42, 43]. Similarly, viruses that do initiate replication
269 within cells may nonetheless undergo abortive infection, due to viral genetic defects, robust
270 cell-intrinsic immunity, or a failure of some segments to be replicated [44-47]. Finally, if many
271 viruses initiate productive infection following transmission, their subsequent loss could be
272 explained by action of innate antiviral immunity or competition among the co-occurring
273 genotypes for limited target cells [48, 49]. Both mechanisms can lead to stochastic rather than
274 selective losses of diversity. For example, a cell that happens to support replication with a
275 relatively short eclipse phase could rapidly seed a single genotype into many secondary cells,
276 allowing it to become dominant.
277
278 In summary, our data reveal a two-stage transmission process in which transfer between hosts
279 can include a large and diverse viral population but early events in the recipient host impose a
280 stringent and stochastic bottleneck. Transmission may therefore represent an opportunity for
281 selection to operate in the earliest phases of infection, before any single genotype sweeps the
282 population. This effect is expected to have a strong influence on viral evolution and is likely
283 relevant across diverse viral families.
284
285

286 MATERIALS & METHODS

287 **Ethical considerations**

288 All the animal experiments were conducted in accordance with the Guide for the Care and Use
289 of Laboratory Animals of the National Institutes of Health. The studies were conducted under
290 animal biosafety level 2 containment and approved by the IACUC of Emory University
291 (PROTO201700595) for the guinea pig (*Cavia porcellus*). The animals were humanely euthanized
292 following guidelines approved by the American Veterinary Medical Association.
293

294 **Cells**

295 Madin–Darby canine kidney (MDCK) cells were a gift from Dr. Daniel Perez, University of
296 Georgia, Athens, GA. A seed stock of MDCK cells at passage 23 was amplified and maintained in
297 Minimal Essential Medium (Gibco) supplemented with 10% fetal bovine serum (FBS; Atlanta
298 Biologicals) and Normocin (Invivogen). 293T cells (ATCC, CRL-3216) were maintained in
299 Dulbecco's Minimal Essential Medium (Gibco) supplemented with 10% FBS and Normocin. All
300 cells were cultured at 37 °C and 5% CO_2 in a humidified incubator. The cell lines were not
301 authenticated. All cell lines were tested monthly for *Mycoplasma* contamination while in use.
302 The medium for the culture of influenza A virus in MDCK cells (virus medium) was prepared by

303 supplementing Minimal Essential Medium with 4.3% bovine serum albumin (BSA; Sigma) and
304 Normocin.
305
306 **Generation of Pan/99 NA-BC plasmid**
307 The region of Pan/99 into which the barcode was inserted was first identified by aligning 593
308 sequences of H3N2-subtype influenza A viruses isolated between 1994 and 2004. A region with
309 many nucleotide substitutions was identified in the NA segment from nucleotide position 484
310 to 532. Twelve synonymous mutations were identified within this region. Double-stranded
311 Ultramers (IDT) were designed that contained degenerate bases with two possible nucleotides
312 at each of the twelve chosen barcode sites
313 (cacagtacatgataggacccccttaycgraccytattgatgaatgarttrggtgtyccattycayytrggracyaagcaagtgtgtatagca
314 tggtcc).
315
316 Site-directed mutagenesis was used to insert an Xho1 restriction site into the wild-type reverse
317 genetics plasmid, pDP Pan/99 NA prior to barcode insertion. This was done to give a means of
318 destroying the parental template following barcode insertion. Successful mutagenesis was
319 confirmed by Xho1 restriction digest and Sanger sequencing. To generate a linearized template
320 for barcode insertion, the following steps were performed. Xho1 digestion of the plasmid stock
321 was followed by phosphatase treatment using rSAP (NEB) to dephosphorylate the cut ends of
322 the plasmid. The plasmid was then amplified by PCR using primers that extend outward from
323 the barcode region: P99_NA_536F 5'-caagtgtgtatagcatggtcc-3' and P99_NA_479R 5'-
324 gggtcctatcatgtactgtg-3'. PCR purification (Qiagen QIAquick PCR Purification Kit) was used to
325 isolate the linearized PCR product, followed by a dual digestion with Dpn1 and Xho1 to remove
326 residual WT plasmid. PCR purification was repeated, and then an assembly reaction using the
327 NEBuilder HiFi DNA Assembly Kit (NEB) was performed to insert the Ultramers into the
328 linearized vector and re-circularize. The product was then transformed into DH5-α cells (NEB).
329 After plating onto LB-amp plates, approximately 1 x 10^4 colonies were collected and pooled into
330 LB-amp culture media and then incubated at 37°C for five hours prior to harvesting the
331 bacterial population for plasmid purification by Maxiprep (Qiagen Plasmid Maxi Kit). The
332 presence of a diverse barcode in the plasmid stock was verified by next generation sequencing
333 (see below). This plasmid stock was then used to generate Pan/99 NA-BC virus by reverse
334 genetics in combination with seven plasmids encoding Pan/99 WT gene segments in a pDP2002
335 vector [50].
336
337 **Viruses**
338 Pan/99 WT and Pan/99 NA-BC viruses were derived from influenza A/Panama/2007/99 (H3N2)
339 virus (Pan/99 WT) and were generated by reverse genetics. In brief, 293T cells transfected with
340 eight reverse genetics plasmids 16–24 h previously were co-cultured with MDCK cells in virus
341 medium at 33°C for 40 h. Recovered virus was propagated in MDCK cells to generate working
342 stocks. Propagation was carried out from low MOI to avoid accumulation of defective viral
343 genomes but with a sufficient viral population size to maintain barcode diversity. Titration of
344 stocks and experimental samples was carried out by plaque assay in MDCK cells. The presence
345 of a diverse barcode in the virus stock was verified by next generation sequencing (see below).

346
347 **Growth kinetics**
348 Replication of Pan/99 WT and Pan/99 NA-BC was determined in triplicate culture wells. MDCK
349 cells in six-well dishes were inoculated at an MOI of 0.01 PFU/cell in PBS. After 1 h incubation at
350 33°C, inoculum was removed, cells were washed 3x with PBS, 2 mL virus medium was added to
351 cells, and dishes were returned to 33°C. A 120 μL volume of culture medium was sampled at
352 the indicated times points and stored at −80 °C. Viral titers were determined by plaque assay on
353 MDCK cells.
354
355 **Guinea pig infections**
356 Female Hartley strain guinea pigs weighing 250–350 g were obtained from Charles River
357 Laboratories and housed by Emory University Department of Animal Resources. Guinea pigs
358 were anesthetized with ketamine (30 mg/kg) and xylazine (4 mg/kg) by intramuscular injection
359 prior to intranasal inoculation or nasal lavage. Virus inoculum was given intranasally in a 300 μl
360 volume of PBS containing 5 x 10^4 PFU of Pan/99 NA-BC. At day 1 post-inoculation, one naïve
361 animal was placed with each inoculated animal in either a cage that allowed for direct physical
362 contact or a cage in which the two animals were separated by a double-walled, perforated
363 metal barrier. Nasal lavage was performed with 1 mL PBS per animal on days 1-7 post-
364 inoculation for inoculated animals and days 2-8 post-inoculation for exposed animals. Collected
365 fluid was divided into aliquots and stored at -80 °C. Viral titers of nasal lavage samples were
366 subsequently determined by plaque assay.
367
368 **Sample processing for next generation sequencing of barcode region**
369 The following method was used to validate plasmid and virus stocks and to evaluate
370 experimental samples. Viral RNA extraction (QiaAmp Viral RNA kit, Qiagen) was performed
371 using a 140 μL volume of each nasal lavage sample or virus stock, followed by reverse
372 transcription (Maxima RT, Thermo Fisher) with pooled Univ.F(A)+6 and Univ.F(G)+6 primers [51,
373 52]. Either cDNA or, for the purpose of plasmid validation, plasmid DNA was subjected to PCR
374 amplification with primers flanking the barcode region. Samples were sent for amplicon
375 sequencing to either Azenta Life Sciences or the Emory National Primate Research Center
376 (ENPRC) Genomics Core. For samples sent to Azenta, PCR to amplify the region containing the
377 barcode was performed using PFU Turbo AD (ThermoFisher) with primers with partial
378 sequencing adapters that generated a 155-nt product (P99_NA_adptr_428-449_F 5'-
379 acactctttccctacacgacgctcttccgatctggaacaacactaaacaacaggc-3' and P99_NA_adptr_563-582_F 5'-
380 gactggagttcagacgtgtgctcttccgatctgcttttccatcgtgacaact-3'). For samples sequenced through
381 ENPRC, amplicon PCR was performed in a similar manner but using primers with sequencing
382 adapters that generated a 100-nt product (P99NA_462F5_adptr 5'-
383 tcgtcggcagcgtcagatgtgtataagagacagggactcctcagtacatgataggacccctt-3' and P99NA_553Rev_adptr
384 5'- gtctcgtgggctcggagatgtgtataagagacagccatgctatacacacttgct-3'). Column-based PCR purification
385 was performed on all samples followed by quantification of DNA using either NanoDrop or
386 Qubit. Samples were normalized to 20 ng/uL in nuclease-free H_20.
387 Samples submitted to Azenta underwent Amplicon-EZ sequencing, an Illumina-based
388 sequencing service compatible with amplicons of 150-500 nt in length that does not include a

389 fragmentation step in library preparation. Amplicons were sequenced as 2 × 250 bp paired
390 reads and demultiplexed prior to delivery. At ENPRC, library preparation was performed with
391 the omission of a tagmentation step and sequencing was performed on a NovaSeq 6000
392 (Illumina) platform. Amplicons were sequenced as 2 x 100 bp paired reads and demultiplexed
393 prior to delivery.
394
395 **Data analysis for amplicon sequencing**
396 Sequences were processed using our custom software, BarcodeID, available in the GitHub
397 repository: https://github.com/Lowen-Lab/BarcodeID. Briefly, BarcodeID uses BBTools [35] to
398 process raw reads, then uses a custom Python script to screen and identify barcode sequences
399 present in each sample, and then calculates diversity statistics and writes summary tables.
400 BBMerge screens reads for adapter sequences and merges forward and reverse reads using
401 default settings, and BBDuk merges reads with a low average quality (<30). BarcodeID then
402 screens each read to verify that the nucleotides at barcode and non-barcode sites match the
403 nucleotides expected at those sites and have sufficient quality (≥35 and ≥25, respectively).
404 Reads with mismatches are excluded from overall barcode sequence counts, but BarcodeID
405 collects all high-quality variant amplicons and calculates overall mismatch rates by site to
406 determine if any mutants with non-barcode alleles are driving any observed barcode dynamics.
407 Samples with evidence of non-barcode driven dynamics were excluded from further analyses.
408 This exclusion criterion was relevant in only one animal, the first aerosol exposed animal within
409 experimental replicate 1. We note that many reads were obtained and analyzed for plaque-
410 positive samples, irrespective of viral titer (**Supplemental Figure 8)**.
411
412 **Preparation of nasal lavage samples for whole-genome sequencing**
413 Viral RNA extraction (QiaAmp Viral RNA kit, Qiagen) was performed using a 140 μL volume of
414 each nasal lavage sample, followed by one-step reverse transcription PCR amplification of full
415 viral genomes using pooled Univ.F(A)+6, Univ.F(G)+6 primers and Univ.R primers and
416 SuperScript III Platinum kit (ThermoFisher) [51, 52]. Following PCR purification (Qiagen
417 QiaQuick PCR Purification Kit), cDNA was processed at the ENPRC core for sequencing on an
418 Illumina NovaSeq 6000 platform. Samples were sequenced as 2 x 100 bp paired reads and
419 demultiplexed prior to delivery.
420
421 **Data analysis for whole-genome sequencing**
422 Whole genome sequencing reads were merged and filtered for low average quality (≥30) using
423 BBMerge, then separated according to the segment using BLAT [53]. The reads were then
424 mapped to their corresponding reference segment using BBMap, with local alignment set to
425 false. From these alignments, we used custom Python scripts to identify iSNVs and reads that
426 map to the barcoded region of the NA segment. Cutoffs for inclusion of iSNVs were set
427 empirically. First, sites were evaluated based on their total coverage and the average quality
428 and mapping statistics. Only sites with ≥100x coverage were considered. For minor variants at
429 these sites to be included in subsequent analyses, they were required to be present at ≥1%
430 frequency and have an average phred score of ≥35, and the reads that contained the minor
431 allele at any given site also had to have sufficient mapping quality to justify inclusion.
432 Specifically, reads containing the minor allele needed an average mapping quality score of ≥40,

433 the average location of the minor allele needed to be ≥20 bases from the nearest end of a read,
434 and the reads overall needed to have ≤2.0 average mismatch and indel counts relative to the
435 reference sequence.
436
437 **Analysis of beta diversity**
438 Dissimilarity between two populations can be measured using beta diversity metrics, and in this
439 study, we used the both Jaccard index and the Bray-Curtis dissimilarity [54, 55]. Both metrics
440 consider the species shared between two populations, and in this case, unique barcodes were
441 considered species. In addition to presence/absence data, the Bray-Curtis dissimilarity also
442 reflects abundance data. For either measure, a value closer to one indicates that the two
443 populations are more dissimilar, whereas a value closer to zero indicates that the two
444 populations are more alike in composition. Pairwise comparisons of barcode data were made
445 between the viral populations present in plaque-positive nasal lavage samples acquired from a
446 contact pair of guinea pigs.
447
448 **Data availability**
449 All data generated or analyzed during this study are included in the manuscript. At the time of
450 initial submission, all raw sequencing data is in the process of being uploaded to NCBI's Sequence
451 Read Archive.
452
453 **Acknowledgments/Funding**
454 This study was supported in part by R01 AI165644 and the National Institute of Allergy and
455 Infectious Diseases (NIAID) Centers of Excellence for Influenza Research and Response (CEIRR)
456 contract no. 75N93021C00017.
457
458 **References**
459 1. Nobusawa E, Sato K. Comparison of the mutation rates of human influenza A and B
460 viruses. J Virol. 2006;80(7):3675-8. doi: 10.1128/JVI.80.7.3675-3678.2006.
461 2. Smith DB, Inglis SC. The mutation rate and variability of eukaryotic viruses: an analytical
462 review. J Gen Virol. 1987;68 (Pt 11):2729-40. doi: 10.1099/0022-1317-68-11-2729.
463 3. Boni MF, Zhou Y, Taubenberger JK, Holmes EC. Homologous recombination is very rare
464 or absent in human influenza A virus. J Virol. 2008;82(10):4807-11. Epub 20080319. doi:
465 10.1128/JVI.02683-07.
466 4. Suarez DL, Senne DA, Banks J, Brown IH, Essen SC, Lee CW, et al. Recombination
467 resulting in virulence shift in avian influenza outbreak, Chile. Emerg Infect Dis. 2004;10(4):693-
468 9. doi: 10.3201/eid1004.030396.
469 5. Marshall N, Priyamvada L, Ende Z, Steel J, Lowen AC. Influenza virus reassortment occurs
470 with high frequency in the absence of segment mismatch. PLoS Pathog. 2013;9(6):e1003421.
471 Epub 20130613. doi: 10.1371/journal.ppat.1003421.
472 6. Barbezange C, Jones L, Blanc H, Isakov O, Celniker G, Enouf V, et al. Seasonal Genetic
473 Drift of Human Influenza A Virus Quasispecies Revealed by Deep Sequencing. Front Microbiol.
474 2018;9:2596. Epub 20181031. doi: 10.3389/fmicb.2018.02596.

475 7. Visher E, Whitefield SE, McCrone JT, Fitzsimmons W, Lauring AS. The Mutational
476 Robustness of Influenza A Virus. PLoS Pathog. 2016;12(8):e1005856. Epub 2016/08/30. doi:
477 10.1371/journal.ppat.1005856.
478 8. Simon B, Pichon M, Valette M, Burfin G, Richard M, Lina B, et al. Whole Genome
479 Sequencing of A(H3N2) Influenza Viruses Reveals Variants Associated with Severity during the
480 2016(-)2017 Season. Viruses. 2019;11(2). Epub 2019/01/31. doi: 10.3390/v11020108.
481 9. Nelson MI, Holmes EC. The evolution of epidemic influenza. Nat Rev Genet.
482 2007;8(3):196-205. Epub 2007/01/31. doi: 10.1038/nrg2053.
483 10. Spielman SJ, Weaver S, Shank SD, Magalis BR, Li M, Kosakovsky Pond SL. Evolution of
484 Viral Genomes: Interplay Between Selection, Recombination, and Other Forces. Methods Mol
485 Biol. 2019;1910:427-68. doi: 10.1007/978-1-4939-9074-0_14.
486 11. Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, et al. Genetic drift,
487 selection and the evolution of the mutation rate. Nat Rev Genet. 2016;17(11):704-14. doi:
488 10.1038/nrg.2016.104.
489 12. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC. The genomic
490 and epidemiological dynamics of human influenza A virus. Nature. 2008;453(7195):615-9. Epub
491 2008/04/18. doi: 10.1038/nature06945.
492 13. Moya A, Holmes EC, Gonzalez-Candelas F. The population genetics and evolutionary
493 epidemiology of RNA viruses. Nat Rev Microbiol. 2004;2(4):279-88. doi: 10.1038/nrmicro863.
494 14. McCrone JT, Lauring AS. Genetic bottlenecks in intraspecies virus transmission. Curr
495 Opin Virol. 2018;28:20-5. Epub 20171103. doi: 10.1016/j.coviro.2017.10.008.
496 15. Valesano AL, Taniuchi M, Fitzsimmons WJ, Islam MO, Ahmed T, Zaman K, et al. The Early
497 Evolution of Oral Poliovirus Vaccine Is Shaped by Strong Positive Selection and Tight
498 Transmission Bottlenecks. Cell Host Microbe. 2021;29(1):32-43 e4. Epub 2020/11/20. doi:
499 10.1016/j.chom.2020.10.011.
500 16. Wang GP, Sherrill-Mix SA, Chang KM, Quince C, Bushman FD. Hepatitis C virus
501 transmission bottlenecks analyzed by deep sequencing. J Virol. 2010;84(12):6218-28. Epub
502 20100407. doi: 10.1128/JVI.02271-09.
503 17. Edwards CT, Holmes EC, Wilson DJ, Viscidi RP, Abrams EJ, Phillips RE, et al. Population
504 genetic estimation of the loss of genetic diversity during horizontal transmission of HIV-1. BMC
505 Evol Biol. 2006;6:28. Epub 20060323. doi: 10.1186/1471-2148-6-28.
506 18. Smith DR, Adams AP, Kenney JL, Wang E, Weaver SC. Venezuelan equine encephalitis
507 virus in the mosquito vector Aedes taeniorhynchus: infection initiated by a small number of
508 susceptible epithelial cells and a population bottleneck. Virology. 2008;372(1):176-86. Epub
509 2007/11/21. doi: 10.1016/j.virol.2007.10.011.
510 19. McCrone JT, Woods RJ, Martin ET, Malosh RE, Monto AS, Lauring AS. Stochastic
511 processes constrain the within and between host evolution of influenza virus. Elife. 2018;7.
512 Epub 2018/04/24. doi: 10.7554/eLife.35962.
513 20. Varble A, Albrecht RA, Backes S, Crumiller M, Bouvier NM, Sachs D, et al. Influenza A
514 virus transmission bottlenecks are defined by infection route and recipient host. Cell Host
515 Microbe. 2014;16(5):691-700. Epub 20141023. doi: 10.1016/j.chom.2014.09.020.
516 21. Frise R, Bradley K, van Doremalen N, Galiano M, Elderfield RA, Stilwell P, et al. Contact
517 transmission of influenza virus between ferrets imposes a looser bottleneck than respiratory

518 droplet transmission allowing propagation of antiviral resistance. Sci Rep. 2016;6:29793. Epub
519 2016/07/20. doi: 10.1038/srep29793.
520 22. Hughes J, Allen RC, Baguelin M, Hampson K, Baillie GJ, Elton D, et al. Transmission of
521 equine influenza virus during an outbreak is characterized by frequent mixed infections and
522 loose transmission bottlenecks. PLoS Pathog. 2012;8(12):e1003081. Epub 20121220. doi:
523 10.1371/journal.ppat.1003081.
524 23. Ferreri LM, Geiger G, Seibert B, Obadan A, Rajao D, Lowen AC, et al. Intra- and inter-host
525 evolution of H9N2 influenza A virus in Japanese quail. Virus Evol. 2022;8(1):veac001. Epub
526 20220108. doi: 10.1093/ve/veac001.
527 24. Braun KM, Haddock Iii LA, Crooks CM, Barry GL, Lalli J, Neumann G, et al. Avian H7N9
528 influenza viruses are evolutionarily constrained by stochastic processes during replication and
529 transmission in mammals. Virus Evol. 2023;9(1):vead004. Epub 20230119. doi:
530 10.1093/ve/vead004.
531 25. Zaraket H, Baranovich T, Kaplan BS, Carter R, Song MS, Paulson JC, et al. Mammalian
532 adaptation of influenza A(H7N9) virus is limited by a narrow genetic bottleneck. Nat Commun.
533 2015;6:6553. Epub 20150408. doi: 10.1038/ncomms7553.
534 26. Wilker PR, Dinis JM, Starrett G, Imai M, Hatta M, Nelson CW, et al. Selection on
535 haemagglutinin imposes a bottleneck during mammalian transmission of reassortant H5N1
536 influenza viruses. Nat Commun. 2013;4:2636. doi: 10.1038/ncomms3636.
537 27. Magurran AE. Ecological Diversity and its Measurement. Princeton, NJ: Princeton
538 University Press; 1988.
539 28. Tao H, Steel J, Lowen AC. Intrahost dynamics of influenza virus reassortment. J Virol.
540 2014;88(13):7485-92. Epub 20140416. doi: 10.1128/JVI.00715-14.
541 29. Xue KS, Bloom JD. Linking influenza virus evolution within and between human hosts.
542 Virus Evol. 2020;6(1):veaa010. Epub 20200217. doi: 10.1093/ve/veaa010.
543 30. Ferguson NM, Galvani AP, Bush RM. Ecological and immunological determinants of
544 influenza evolution. Nature. 2003;422(6930):428-33. doi: 10.1038/nature01509.
545 31. Koelle K, Cobey S, Grenfell B, Pascual M. Epochal evolution shapes the phylodynamics of
546 interpandemic influenza A (H3N2) in humans. Science. 2006;314(5807):1898-903. doi:
547 10.1126/science.1132745.
548 32. Fitch WM, Leiter JM, Li XQ, Palese P. Positive Darwinian evolution in human influenza A
549 viruses. Proc Natl Acad Sci U S A. 1991;88(10):4270-4. doi: 10.1073/pnas.88.10.4270.
550 33. Debbink K, McCrone JT, Petrie JG, Truscon R, Johnson E, Mantlo EK, et al. Vaccination
551 has minimal impact on the intrahost diversity of H3N2 influenza viruses. PLoS Pathog.
552 2017;13(1):e1006194. Epub 20170131. doi: 10.1371/journal.ppat.1006194.
553 34. Han AX, Felix Garza ZC, Welkers MR, Vigeveno RM, Tran ND, Le TQM, et al. Within-host
554 evolutionary dynamics of seasonal and pandemic human influenza A viruses in young children.
555 Elife. 2021;10. Epub 20210803. doi: 10.7554/eLife.68917.
556 35. Valesano AL, Fitzsimmons WJ, McCrone JT, Petrie JG, Monto AS, Martin ET, et al.
557 Influenza B Viruses Exhibit Lower Within-Host Diversity than Influenza A Viruses in Human
558 Hosts. J Virol. 2020;94(5). Epub 20200214. doi: 10.1128/JVI.01710-19.
559 36. Valesano AL, Rumfelt KE, Dimcheff DE, Blair CN, Fitzsimmons WJ, Petrie JG, et al.
560 Temporal dynamics of SARS-CoV-2 mutation accumulation within and across infected hosts.
561 PLoS Pathog. 2021;17(4):e1009499. Epub 20210407. doi: 10.1371/journal.ppat.1009499.

562 37. Gu H, Quadeer AA, Krishnan P, Ng DYM, Chang LDJ, Liu GYZ, et al. Within-host genetic
563 diversity of SARS-CoV-2 lineages in unvaccinated and vaccinated individuals. Nat Commun.
564 2023;14(1):1793. Epub 20230331. doi: 10.1038/s41467-023-37468-y.
565 38. Braun KM, Moreno GK, Wagner C, Accola MA, Rehrauer WM, Baker DA, et al. Acute
566 SARS-CoV-2 infections harbor limited within-host diversity and transmit via tight transmission
567 bottlenecks. PLoS Pathog. 2021;17(8):e1009849. Epub 20210823. doi:
568 10.1371/journal.ppat.1009849.
569 39. Kistler KE, Huddleston J, Bedford T. Rapid and parallel adaptive mutations in spike S1
570 drive clade success in SARS-CoV-2. Cell Host Microbe. 2022;30(4):545-55 e4. Epub 20220322.
571 doi: 10.1016/j.chom.2022.03.018.
572 40. Tellier R. Review of aerosol transmission of influenza A virus. Emerg Infect Dis.
573 2006;12(11):1657-62. doi: 10.3201/eid1211.060426.
574 41. Le Sage V, Jones JE, Kormuth KA, Fitzsimmons WJ, Nturibi E, Padovani GH, et al. Pre-
575 existing heterosubtypic immunity provides a barrier to airborne transmission of influenza
576 viruses. PLoS Pathog. 2021;17(2):e1009273. Epub 20210218. doi:
577 10.1371/journal.ppat.1009273.
578 42. Kaler L, Iverson E, Bader S, Song D, Scull MA, Duncan GA. Influenza A virus diffusion
579 through mucus gel networks. Commun Biol. 2022;5(1):249. Epub 20220322. doi:
580 10.1038/s42003-022-03204-3.
581 43. Zanin M, Baviskar P, Webster R, Webby R. The Interaction between Respiratory
582 Pathogens and Mucus. Cell Host Microbe. 2016;19(2):159-68. doi: 10.1016/j.chom.2016.01.001.
583 44. Jacobs NT, Onuoha NO, Antia A, Steel J, Antia R, Lowen AC. Incomplete influenza A virus
584 genomes occur frequently but are readily complemented during localized viral spread. Nat
585 Commun. 2019;10(1):3526. Epub 20190806. doi: 10.1038/s41467-019-11428-x.
586 45. Brooke CB. Biological activities of 'noninfectious' influenza A virus particles. Future Virol.
587 2014;9(1):41-51. doi: 10.2217/fvl.13.118.
588 46. Brooke CB, Ince WL, Wrammert J, Ahmed R, Wilson PC, Bennink JR, et al. Most influenza
589 a virions fail to express at least one essential viral protein. J Virol. 2013;87(6):3155-62. Epub
590 20130102. doi: 10.1128/JVI.02284-12.
591 47. Marcus PI, Ngunjiri JM, Sekellick MJ. Dynamics of biologically active subpopulations of
592 influenza virus: plaque-forming, noninfectious cell-killing, and defective interfering particles. J
593 Virol. 2009;83(16):8122-30. Epub 20090603. doi: 10.1128/JVI.02680-08.
594 48. Pearson JE, Krapivsky P, Perelson AS. Stochastic theory of early viral infection:
595 continuous versus burst production of virions. PLoS Comput Biol. 2011;7(2):e1001058. Epub
596 20110203. doi: 10.1371/journal.pcbi.1001058.
597 49. Ramos I, Smith G, Ruf-Zamojski F, Martinez-Romero C, Fribourg M, Carbajal EA, et al.
598 Innate Immune Response to Influenza Virus at Single-Cell Resolution in Human Epithelial Cells
599 Revealed Paracrine Induction of Interferon Lambda 1. J Virol. 2019;93(20). Epub 20190930. doi:
600 10.1128/JVI.00559-19.
601 50. Chen H, Ye J, Xu K, Angel M, Shao H, Ferrero A, et al. Partial and full PCR-based reverse
602 genetics strategy for influenza viruses. PLoS One. 2012;7(9):e46378. Epub 20120928. doi:
603 10.1371/journal.pone.0046378.

604 51. Delima GK, Ganti K, Holmes KE, Shartouny JR, Lowen AC. Influenza A virus coinfection
605 dynamics are shaped by distinct virus-virus interactions within and between cells. PLoS Pathog.
606 2023;19(3):e1010978. Epub 20230302. doi: 10.1371/journal.ppat.1010978.
607 52. Zhou B, Wentworth DE. Influenza A virus molecular virology techniques. Methods Mol
608 Biol. 2012;865:175-92. Epub 2012/04/25. doi: 10.1007/978-1-61779-621-0_11.
609 53. Kent WJ. BLAT--the BLAST-like alignment tool. Genome Res. 2002;12(4):656-64. doi:
610 10.1101/gr.229202.
611 54. Chung NC, Miasojedow B, Startek M, Gambin A. Jaccard/Tanimoto similarity test and
612 estimation methods for biological presence-absence data. BMC Bioinformatics. 2019;20(Suppl
613 15):644. Epub 20191224. doi: 10.1186/s12859-019-3118-5.
614 55. J. Roger Bray JTC. An Ordination of the Upland Forest Communities of Southern
615 Wisconsin. Ecological Monographs. 1957;27(4):325-49. doi: 10.2307/1942268.
616

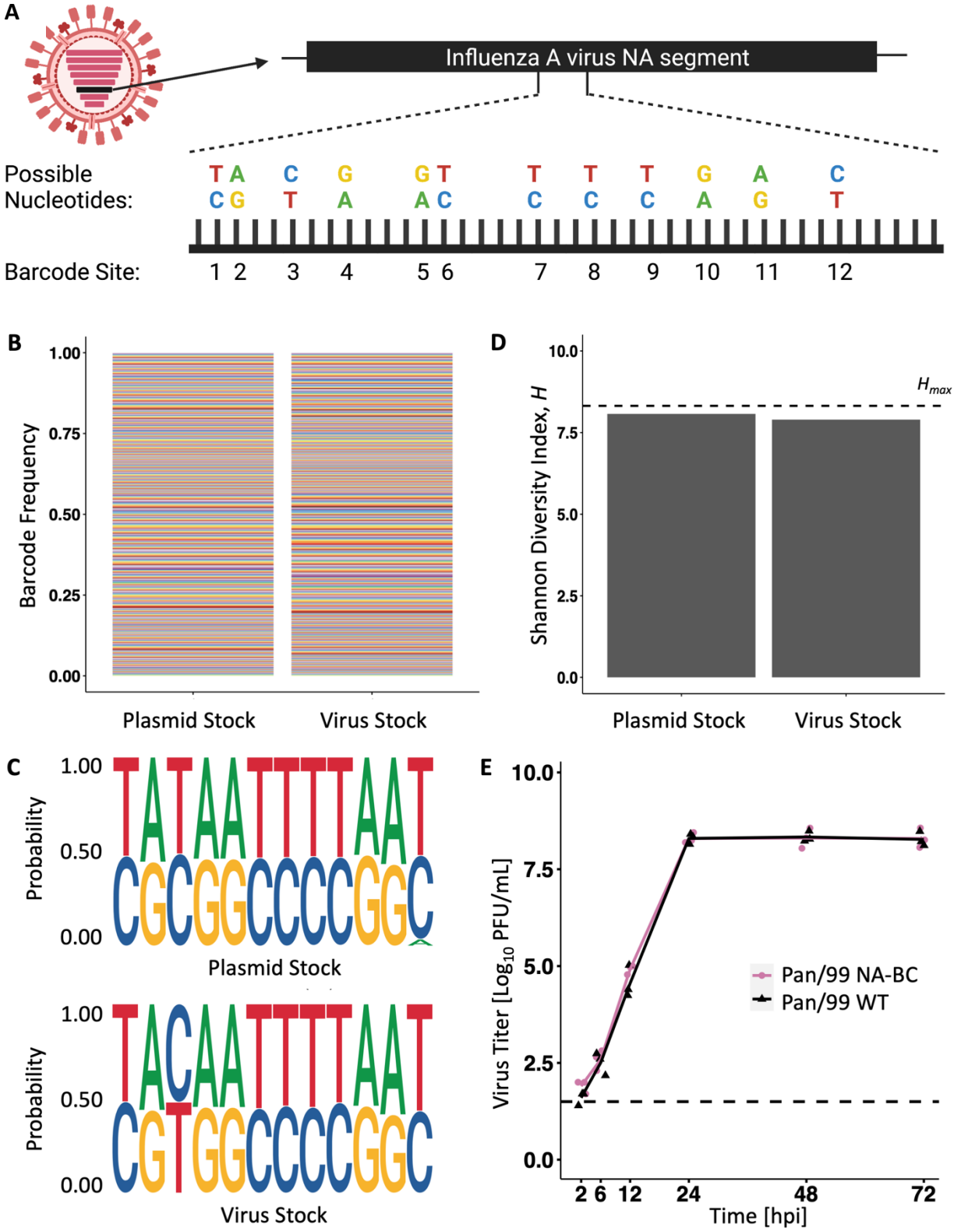


Figure 1. NA barcode diversity is maintained in both plasmid and virus stocks, and the barcode does not affect overall fitness. A) Barcode design for the NA segment of influenza

620 A/Panama/2007/99 (H3N2) virus. An alignment of 593 NA sequences derived from H3N2 subtype
621 human influenza A viruses circulating from 1994 to 2004 showed a cluster of single nucleotide
622 variants occurring in the region from nucleotides 484 to 534. Twelve polymorphic sites containing
623 synonymous SNVs within this region were selected for the barcode. At each of the twelve sites,
624 one of two nucleotides is possible, allowing for up to 4096 unique barcodes within the
625 population. **B)** Barcodes detected in the pDP Pan/99 NA-BC plasmid preparation and passage 1
626 stock of Pan/99 NA-BC virus. Colors represent unique barcodes, and their frequencies within the
627 stock are indicated by the height of the color. **C)** Sequence logo plots demonstrate the sequence
628 motifs present in plasmid and passage 1 virus stocks. Each pair of nucleotides represents one of
629 the twelve bi-allelic sites. The height of the letter indicates the corresponding nucleotide
630 frequency in the stock. **D)** Shannon Diversity Index (*H*) of the stock samples is compared to the
631 maximum possible diversity for this system (H_{max} = 8.32, shown with a horizontal dashed line).
632 This theoretical maximum reflects a population in which all 4096 potential barcodes are equally
633 represented. **E)** Multicycle replication of Pan/99 WT (black) and Pan/99 NA-BC (pink) viruses
634 revealed similar growth kinetics. Infections were performed in triplicate and individual data
635 points show results from individual cell culture dishes. Lines connect mean titers at each time
636 point. Dashed line indicates limit of detection for plaque assay (50 PFU/mL). Two-way ANOVA
637 showed no significant difference between viruses at any time point (p = 0.11). Figure 1A created
638 in BioRender.
639
640

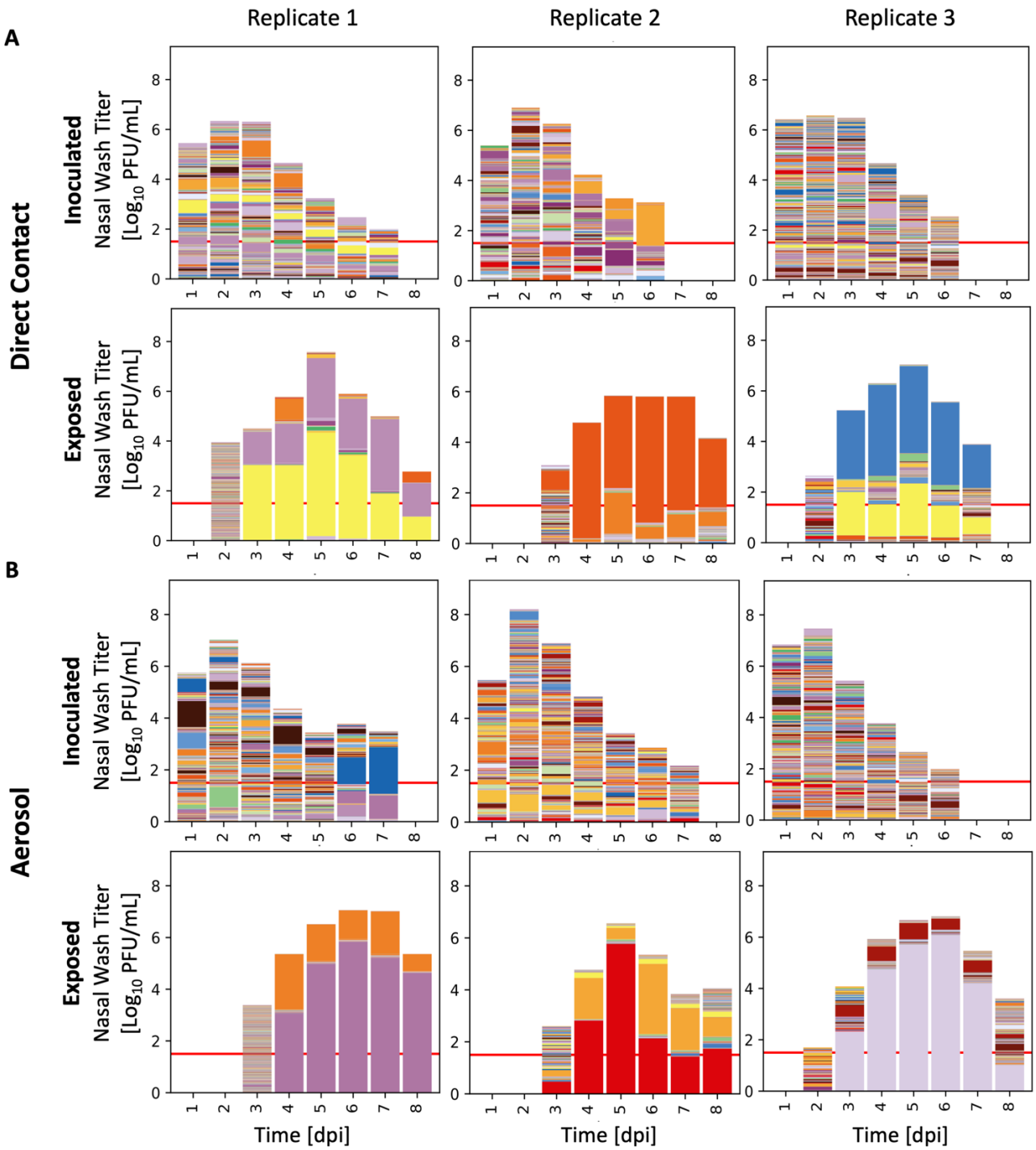


Figure 2. Population diversity declines between inoculated and exposed guinea pigs. Colors in the stacked bar plots (**A, B**) represent unique barcodes, and the height of each color indicates the relative frequency within the sample. Nasal lavage titers are indicated by the height of the bar. Red lines show the limit of detection of the plaque assay (50 PFU/mL). Plots for individual animals are paired with those of their cage mate. Representative pairs for direct contact (**A**) and aerosol (**B**) exposure are shown from three experimental replicates. Data from all transmission pairs (eight per replicate) are shown in Supplemental Figure 2.

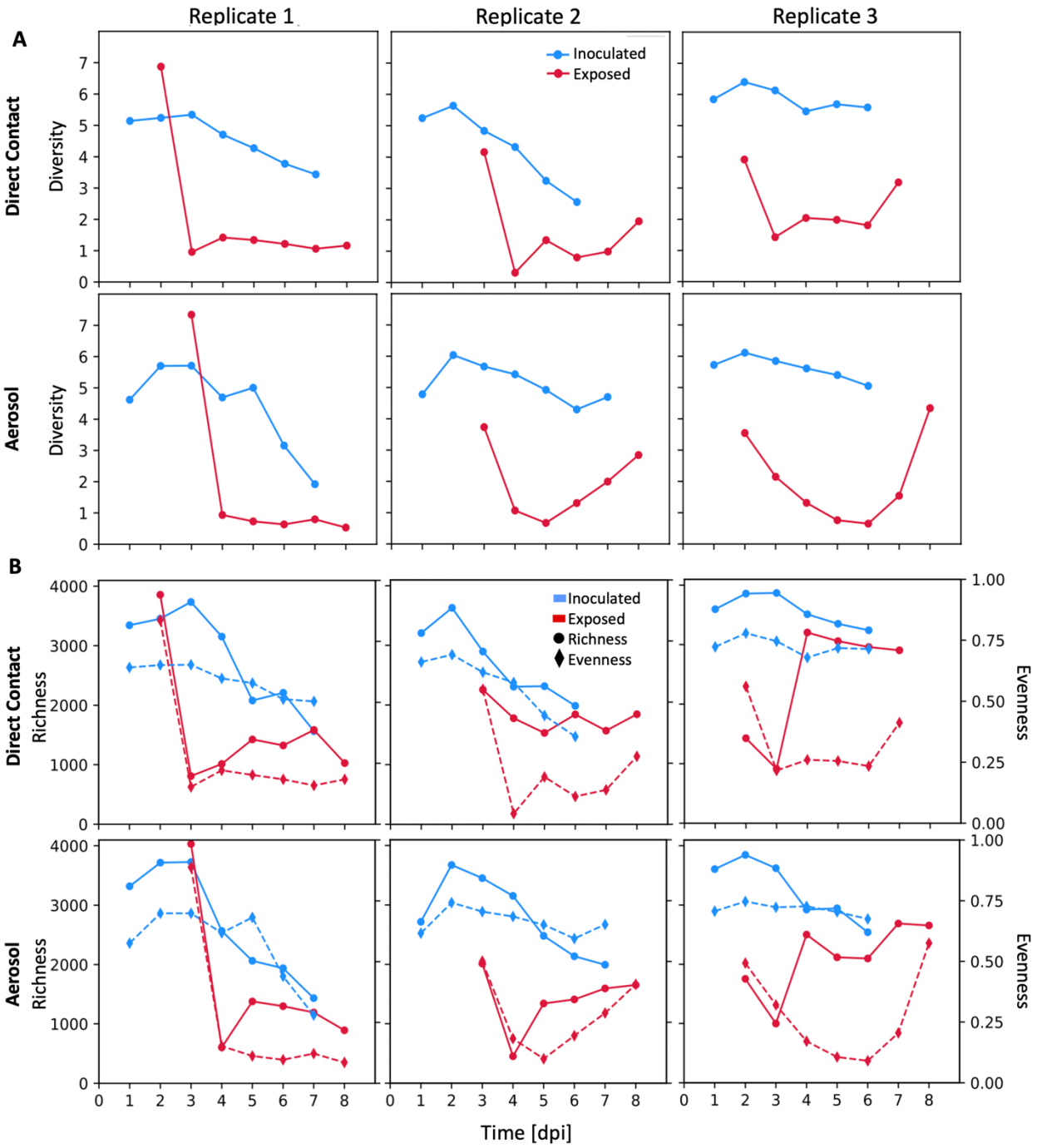


Figure 3. Initial high viral diversity in exposed animals plummets after the first 1-2 days of viral positivity. The Shannon Diversity index (**A**), Chao richness (**B**, right axis), and evenness (**B**, left axis) are shown for viral populations present in representative animals from each replicate (blue = inoculated animals, red = exposed, solid line = Chao richness, dashed line = evenness). (**A,B**) Top row shows direct contact animal pairs, and bottom row shows aerosol transmission pairs. Data from all transmission pairs (eight per replicate) are shown in Supplemental Figure 3.

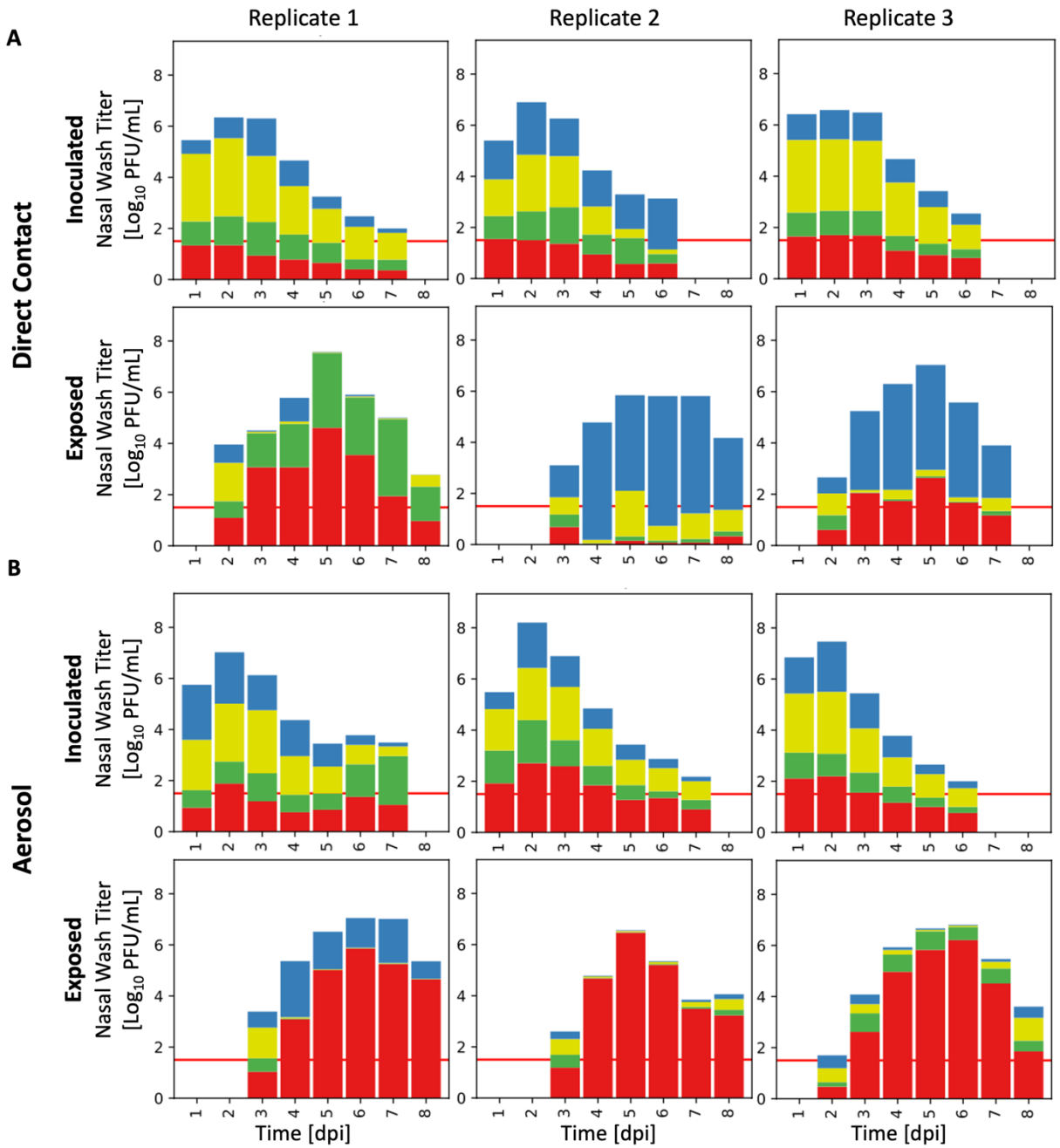
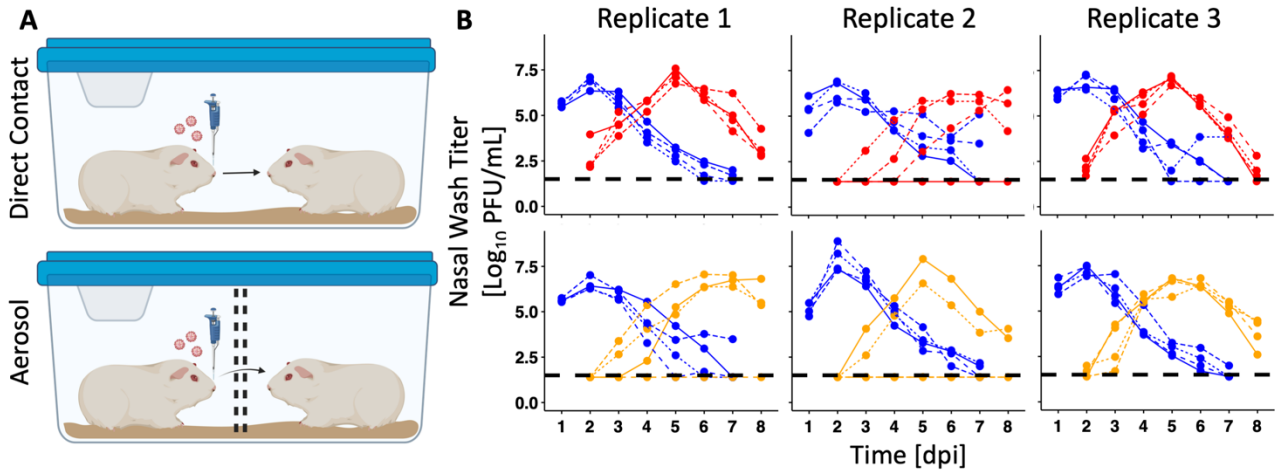
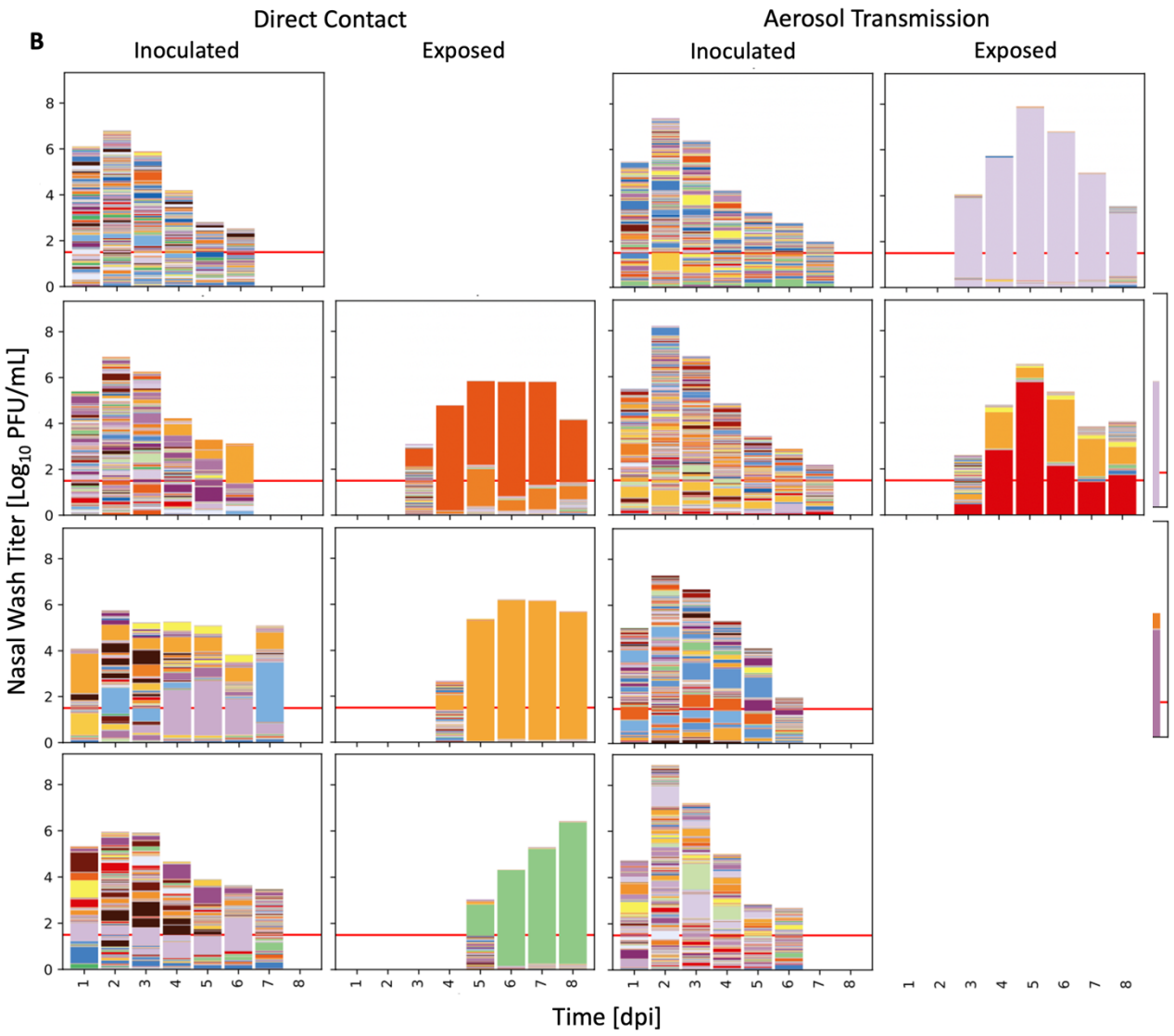


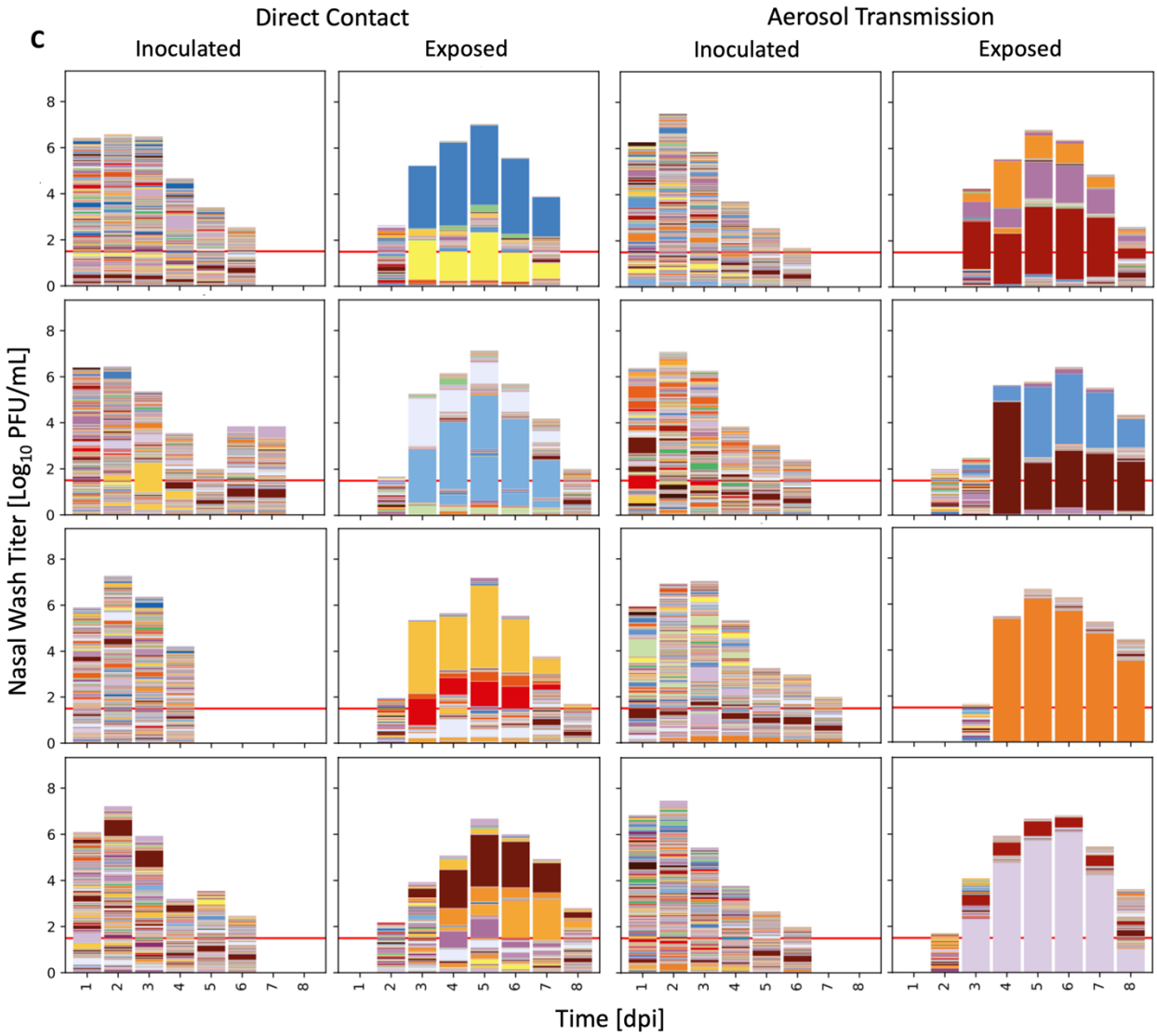
Figure 4. When barcode data are simplified into just four haplotypes, population composition remains consistent in inoculated animals and stochastic effects remain pronounced in exposed animals. Barcodes were binned according to the first two barcode sites for a total of four possible barcodes. Colors in the stacked bar plots (**A, B**) represent unique barcodes, and the height of each color indicates the relative frequency within the sample. Nasal lavage titers are indicated by the height of the bar. Red lines show LOD of plaque assays (50 PFU/mL). Plots for individual animals are paired with those of their cage mate. Representative pairs for direct contact (**A**) and aerosol (**B**) exposure are shown from three experimental replicates. Data for all transmission pairs are shown in Supplemental Figure 5.



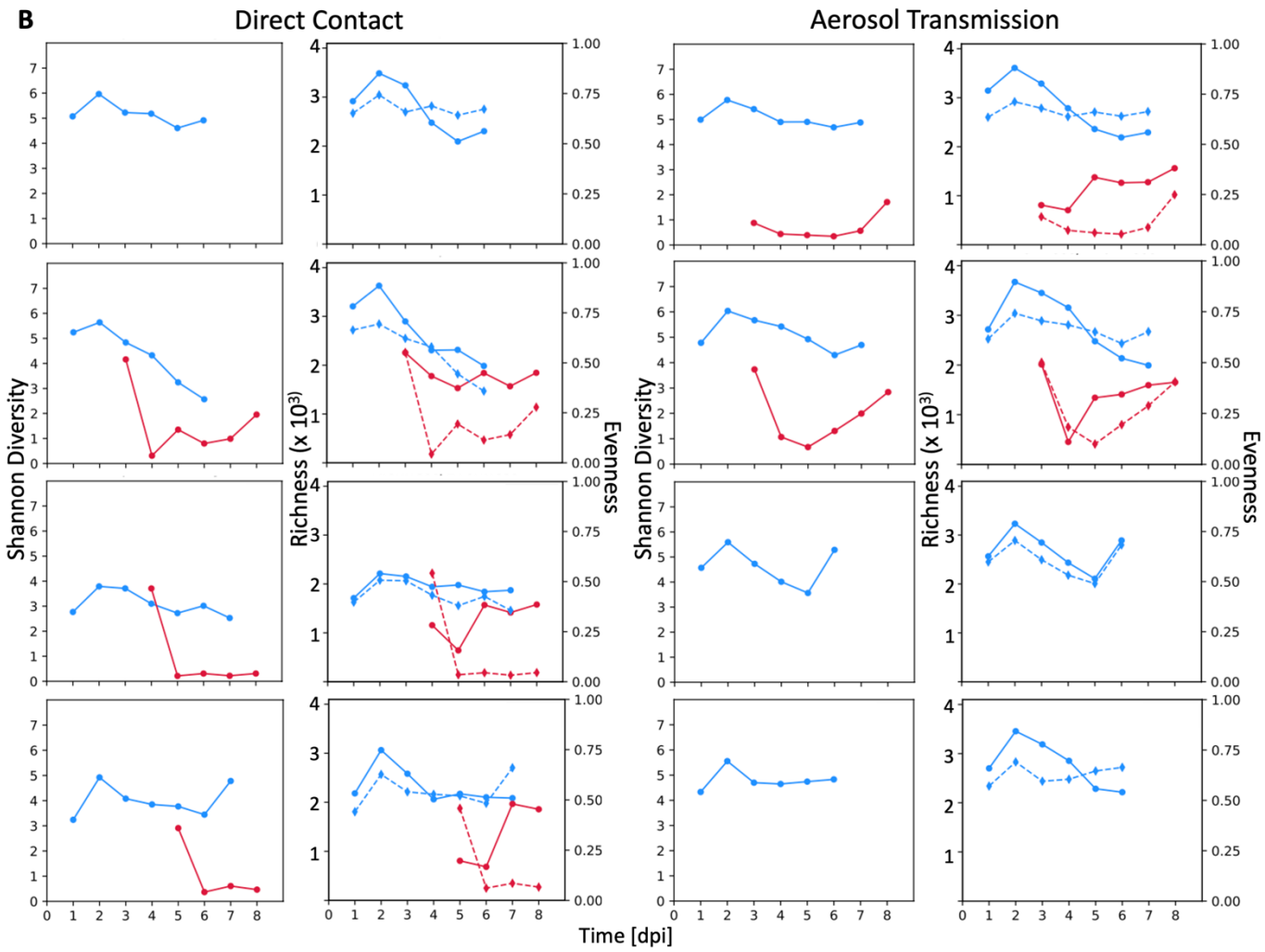
Supplemental Figure 1. Pan/99 NA-BC virus infects inoculated animals and transmits to exposed animals. A) Schematic showing experimental set-up for transmission experiments in guinea pigs. For each of the three experimental replicates performed, eight guinea pigs were intranasally inoculated with 5 x 10^4 PFU of Pan/99 NA-BC in 300 **µ**L PBS. Twenty-four hours post-inoculation, a single naïve animal was placed with each inoculated animal in cages that allowed for animals to directly contact one another (n = 4) or that separated the animals by a double-walled, perforated metal barrier (n = 4). **B)** Viral titers from nasal lavage samples in direct contact (top) and aerosol (bottom) exposure settings from replicates 1, 2, and 3. Blue lines represent inoculated animals; red or yellow lines indicate exposed animals. The dashed black line represents the limit of detection of plaque assays (50 PFU/mL). Paired animals share the same line type. Supplemental Figure 1A created in BioRender.


B
Direct Contact
Aerosol Transmission
Inoculated
Exposed
Inoculated
Exposed
Nasal Wash Titer [Log10 PFU/mL]
Time [dpi]

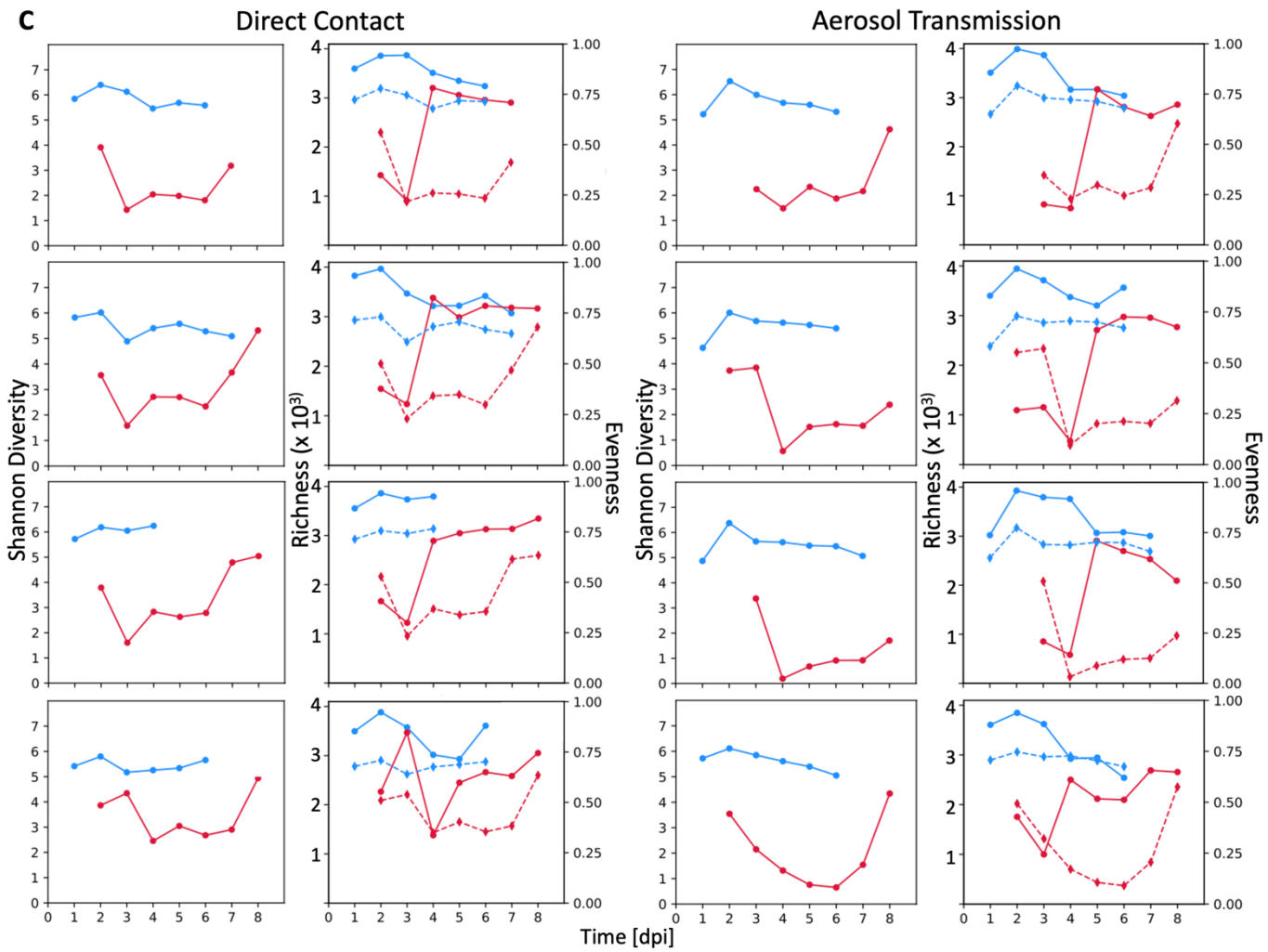
678



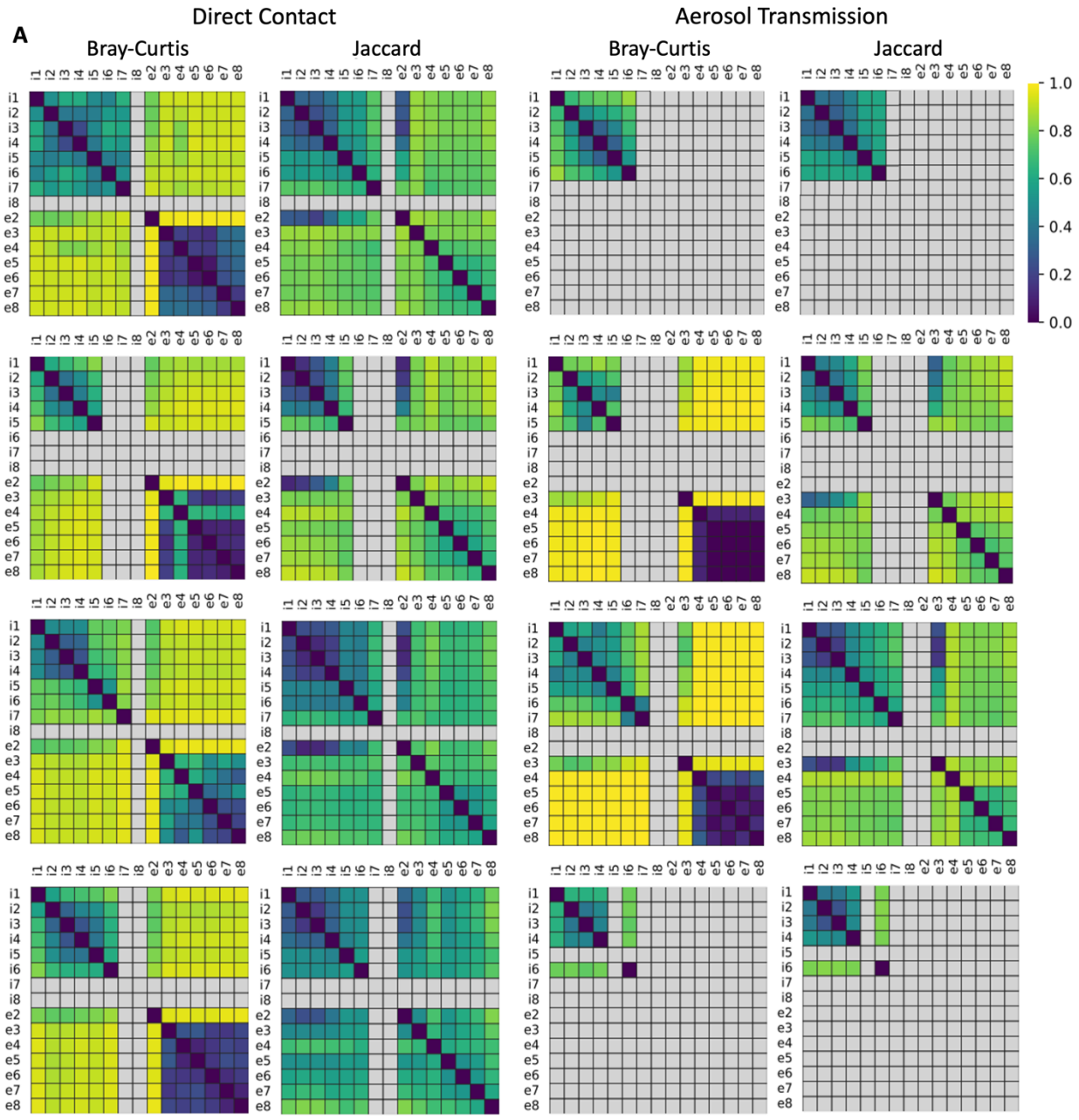
Supplemental Figure 2. Tracking viral barcode diversity within inoculated animals and through transmission to exposed animals. Data relate to that of Figure 2 and show relative barcode frequencies for all transmission pairs across replicates 1, 2, and 3 **(A, B**, and **C**, respectively). Only samples that were plaque-positive are shown. Colors in the stacked bar plots represent unique barcodes, and the height of each color indicates the relative frequency within the sample. Nasal lavage titers are indicated by the height of the bar. Red lines show LOD of plaque assays. Plots for individual animals are paired with those of their cage mate. For the exposed animal in the first aerosol transmission pair of Replicate 1, most reads were discarded, and the data from this animal were excluded from further analyses. The lack of data for other exposed animals indicates a lack of transmission.


B
Direct Contact
Aerosol Transmission
Shannon Diversity
Richness (x 10³)
Evenness
Time [dpi]

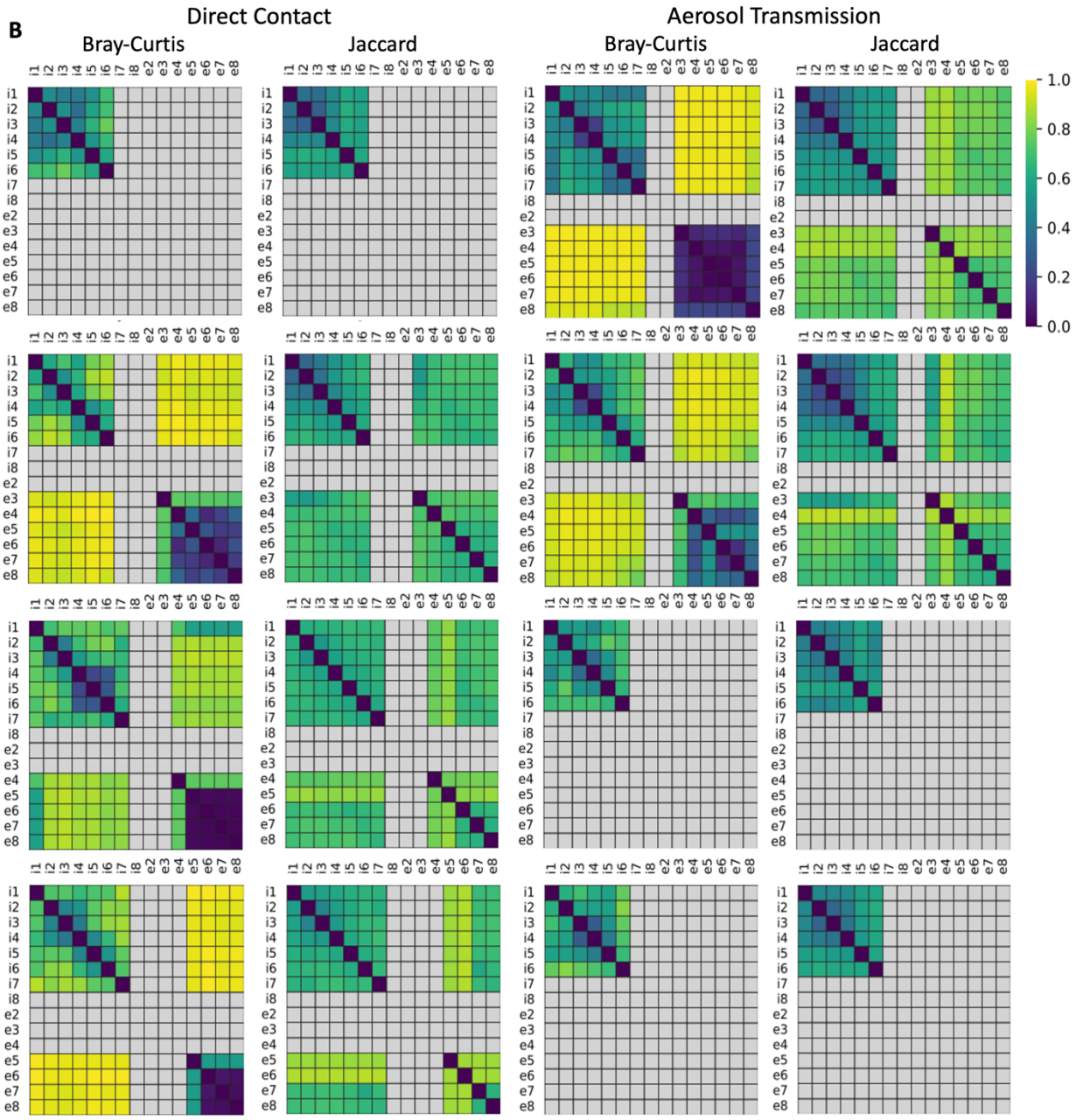
691



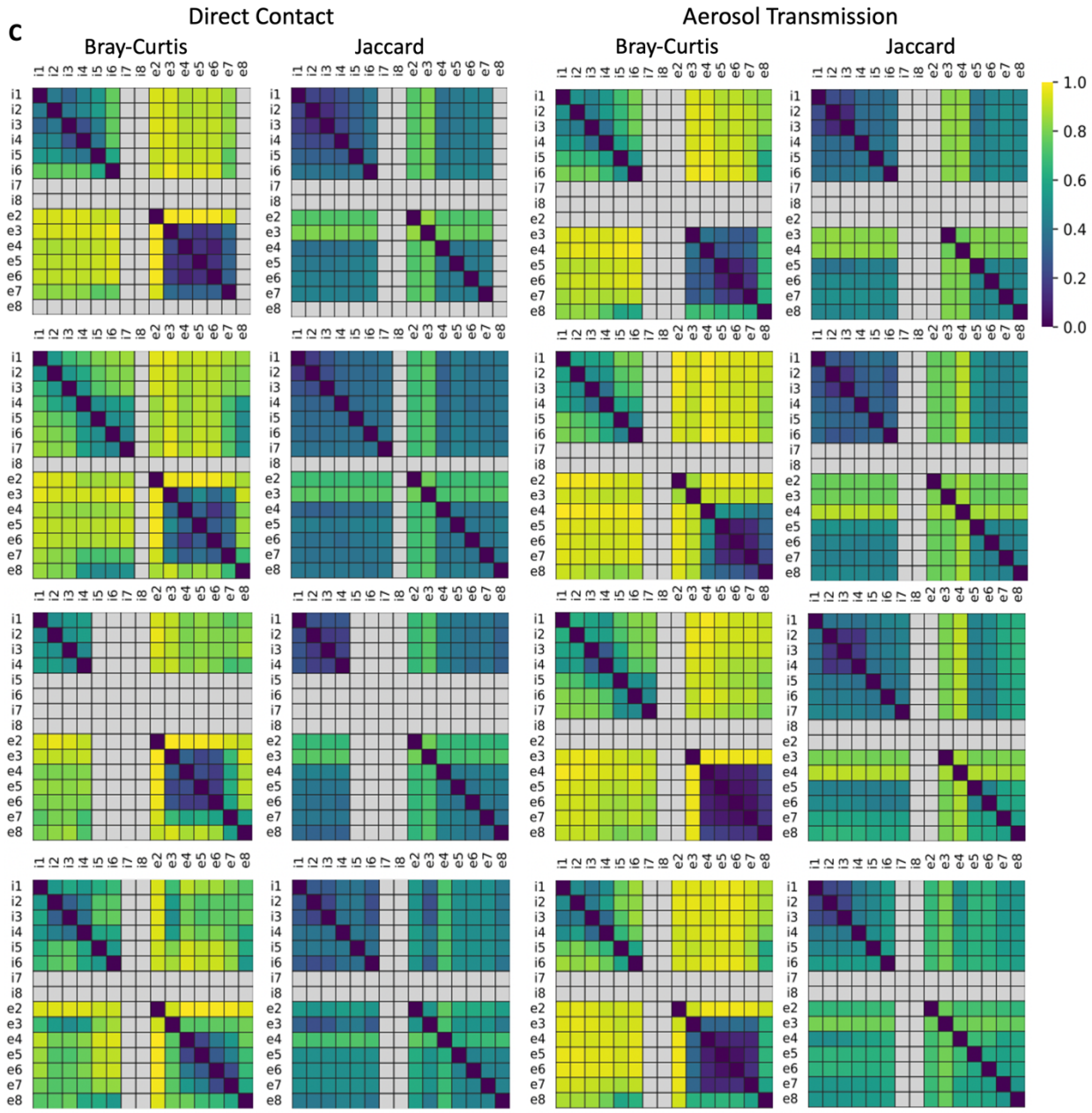
692 **Supplemental Figure 3. Changes in richness and evenness both contribute to alterations in**
693 **diversity.** Data relate to that of Figure 3. Shannon Diversity (left), Chao richness (right, left axis),
694 and evenness (right, right axis) were determined for inoculated animals (dpi 1-7) and exposed
695 animals (dpi 2-8) in replicates 1, 2, and 3 (**A, B,** and **C**, respectively). Data from cage mates are
696 plotted on the same facet (blue = inoculated animals, red = exposed, solid line = Chao richness,
697 dashed line = evenness).


A
Direct Contact
Aerosol Transmission
Bray-Curtis
Jaccard
Bray-Curtis
Jaccard
1.0
0.8
0.6
0.4
0.2
0.0

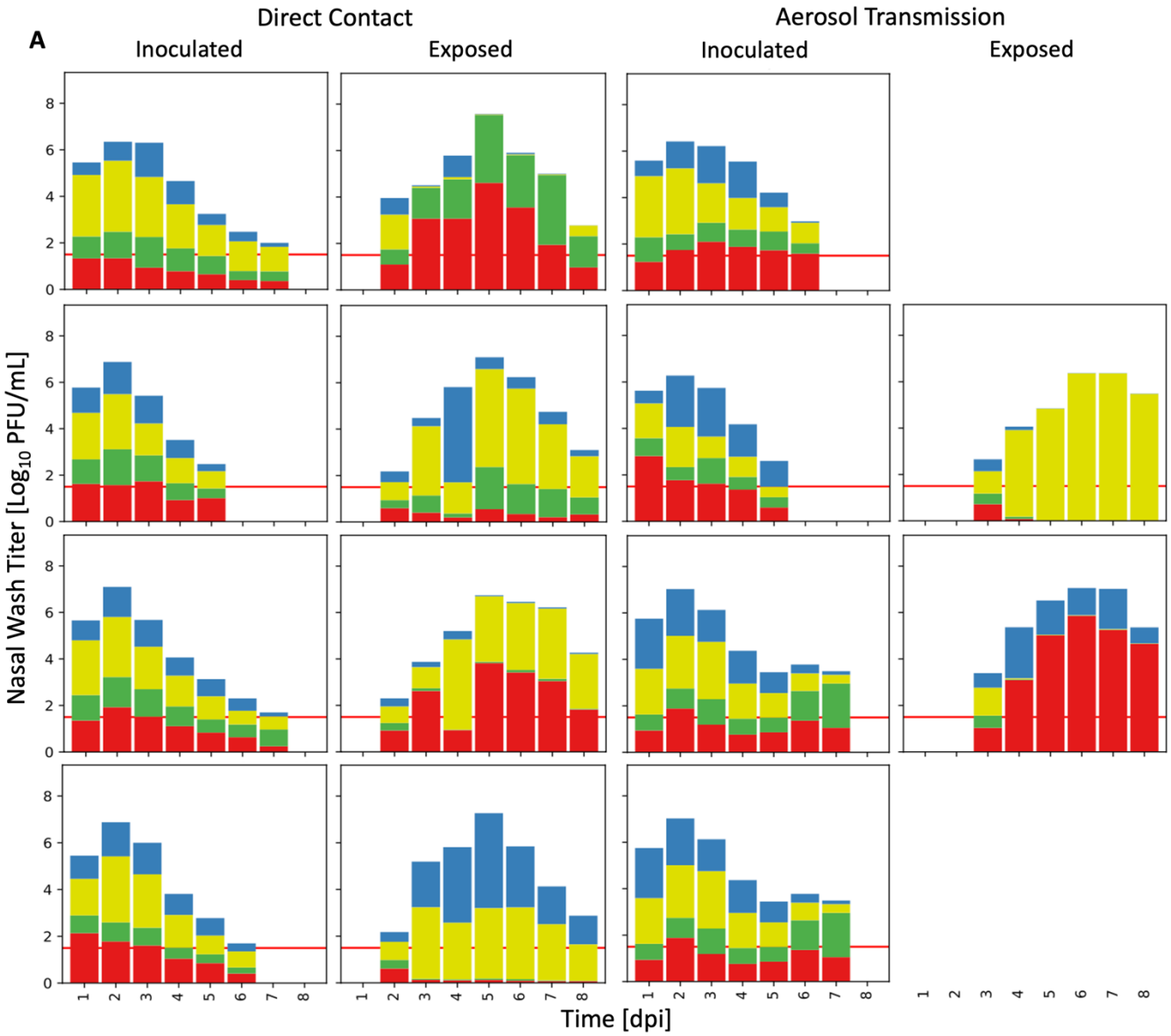
698


B
Direct Contact
Aerosol Transmission
Bray-Curtis
Jaccard
Bray-Curtis
Jaccard
i1 i2 i3 i4 i5 i6 i7 i8 e2 e3 e4 e5 e6 e7 e8
1.0
0.8
0.6
0.4
0.2
0.0

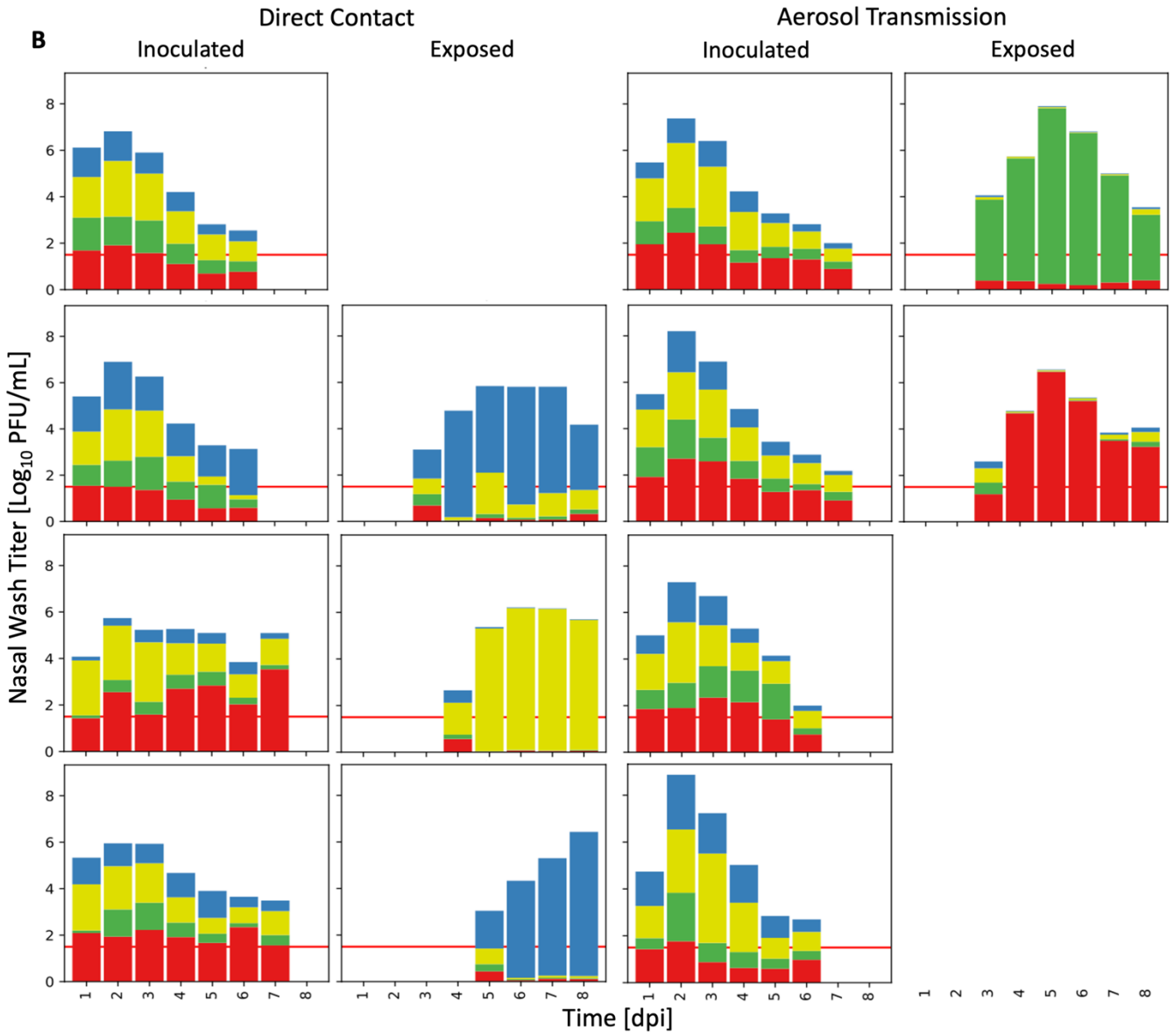
699



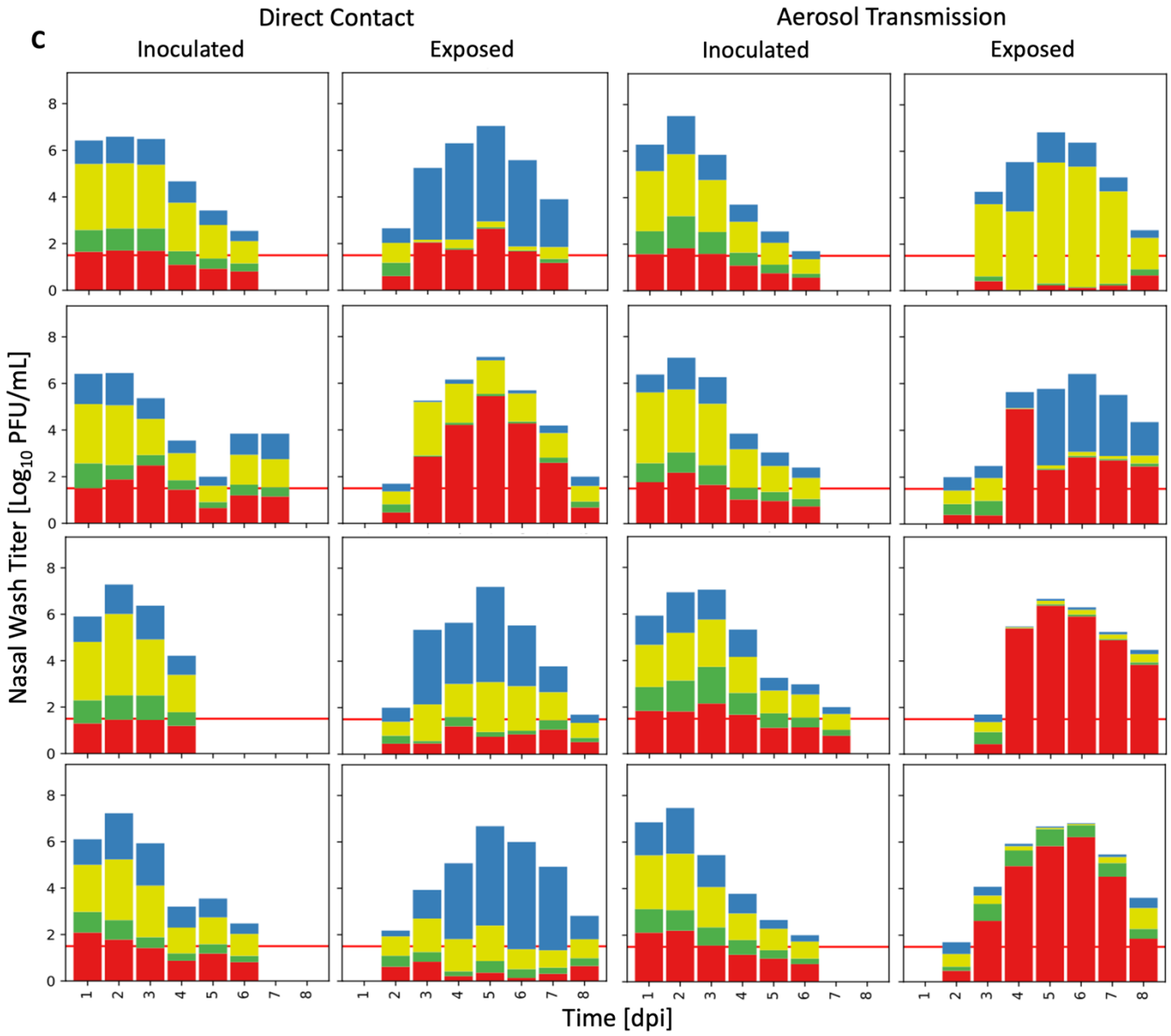
Supplemental Figure 4. Pairwise matrices of dissimilarity demonstrate similar population composition among individual animals but not between paired animals. We calculated indices of dissimilarity (Bray-Curtis and Jaccard) between samples taken on different days from a given animal and between samples collected from paired inoculated and exposed animals. Panels A, B and C show results from experimental replicates 1, 2 and 3, respectively. Axes labels indicate whether an animal was inoculated (*i*) or exposed (*e*), with the accompanying number specifying the day post-inoculation. Warmer colors (more yellow) indicate greater dissimilarity between samples, while cooler colors (more purple) indicate greater similarity between samples. Gray squares signify samples that were negative or omitted due to low sequencing coverage.


A
Direct Contact
Aerosol Transmission
Inoculated
Exposed
Inoculated
Exposed
Nasal Wash Titer [Log10 PFU/mL]
Time [dpi]

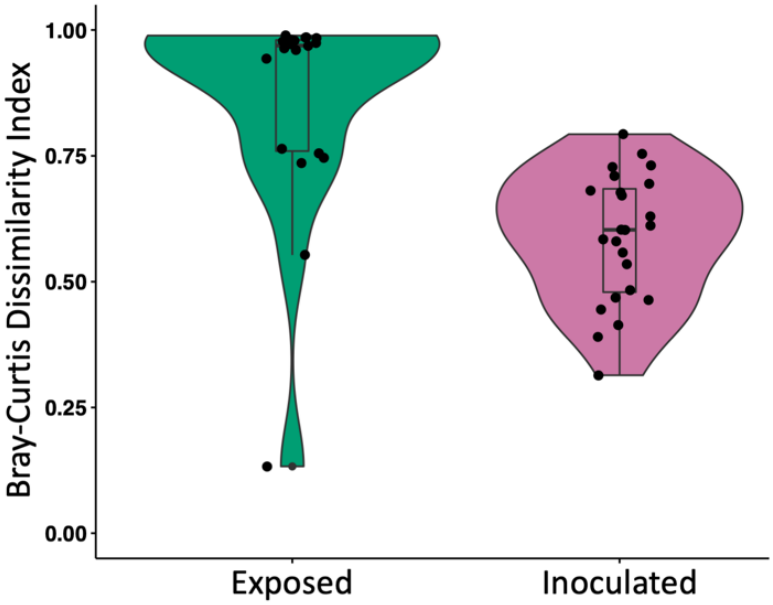
711


B
Direct Contact
Aerosol Transmission
Inoculated
Exposed
Inoculated
Exposed
Nasal Wash Titer [Log10 PFU/mL]
Time [dpi]

712

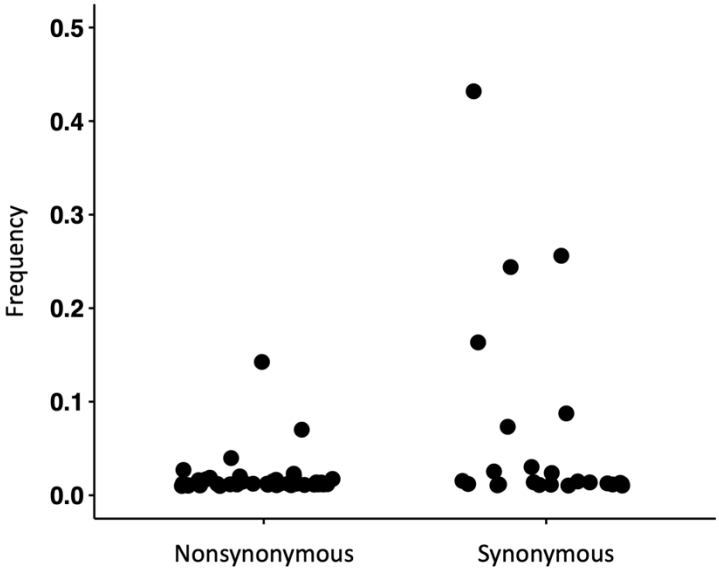


713 **Supplemental Figure 5. Simplification of data into four haplotypes.** Data relate to that of Figure
714 4 and show stacked bar plots with 4096 barcodes divided into 4 groups. Data from paired
715 inoculated and exposed animals are side-by-side and arranged by experimental replicates 1, 2,
716 and 3 (**A, B,** and **C**, respectively) and type of transmission. Arrangement corresponds to that of
717 Supplemental Figure 2. Only samples that were positive by plaque assay are shown. Colors in the
718 stacked bar represent the four haplotype classes into which barcodes were divided based on the
719 nucleotide identity of the first two sites. The height of each color indicates the relative frequency
720 within the sample. Nasal lavage titers are indicated by the height of the bar. Red lines show limit
721 of detection of plaque assays. For the exposed animal in the first aerosol transmission pair of
722 Replicate 1, most reads were discarded, and the data from this animal were excluded from
723 further analyses.



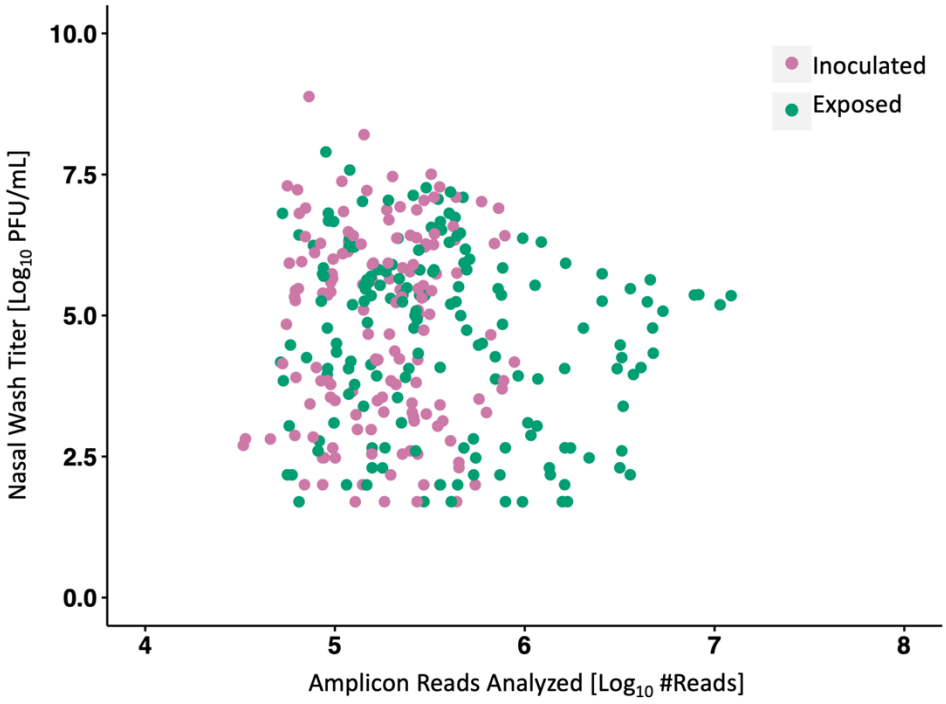
Supplemental Figure 6. Barcode composition in exposed animals changes significantly between the first and subsequent days of detection. Pairwise Bray-Curtis indices for nasal lavage samples from exposed animals on the first versus the third days of positivity ("Exposed", green) and for inoculated animals on the first versus the third days of positivity ("Inoculated", pink) show a significant difference (Mann Whitney U Test, p = 0.0000090), indicating greater dissimilarity between first and subsequent days in exposed animals. Data are presented as violin plots featuring a box plot. The bounds of the box show the first and third quartiles. Whiskers indicate 1.5 times the interquartile range. The bounds of the violin plots indicate the minima and maxima of the specified data.

750
751



752
753
754 **Supplemental Figure 7. Most iSNVs within exposed animals are rare.** The frequencies of
755 nonsynonymous and synonymous iSNVs detected on the first or second day of positivity in
756 exposed animals are shown for experimental replicates 1 and 2. Data from all 16 animals and
757 both time points are plotted together. (This analysis was not performed for replicate 3). No iSNVs
758 were found on multiple days above the frequency cutoff of 1%.
759

760



779

780 **Supplemental Figure 8. Many reads were obtained and analyzed for plaque-positive samples,**
781 **irrespective of viral titer.** Data from inoculated animals are shown in pink and those from
782 exposed animals in green.