Collective interactions augment influenza A virus replication in a host-dependent manner

Kara L. Phipps[a], Ketaki Ganti[a], Silvia Carnaccini[b], Miglena Manandhar[c], Nathan T. Jacobs[a], Brett E. Pickett[c], Gene S. Tan[c,d,f] Lucas M. Ferreri [b], Daniel R. Perez [b,e] and Anice C. Lowen[a,d]*

[a]Department of Microbiology and Immunology, Emory University School of Medicine, Atlanta, Georgia, United States of America

[b]Department of Population Health, Poultry Diagnostic and Research Center, University of Georgia, Athens, Georgia, United States of America.

[c]J. Craig Venter Institute, La Jolla, California, United States of America

[d]Emory-UGA Center of Excellence for Influenza Research and Surveillance (CEIRS), Atlanta, Georgia, United States of America

[e]Center of Research on Influenza Pathogenesis (CRIP) CEIRS, New York, New York, United States of America.

[f]Division of Infectious Disease, Department of Medicine, University of California San Diego, La Jolla, CA, United States of America

*Address correspondence: anice.lowen@emory.edu

Abstract

Infection with a single influenza A virus (IAV) is only rarely sufficient to initiate productive infection. Here, we exploit both single-cell approaches and whole-animal systems to show that the extent of IAV reliance on multiple infection varies with virus strain and host species. Influenza A/guinea fowl/HK/WF10/99 (H9N2) [GFHK99] virus exhibits strong dependence on collective interactions in mammalian systems. This reliance focuses viral progeny production within coinfected cells and therefore results in frequent genetic exchange through reassortment. In contrast, GFHK99 virus has greatly reduced dependence on multiple infection in avian systems, indicating a role for host factors in viral collective interactions. Genetic mapping implicated the viral polymerase as a major driver of multiple infection dependence. Mechanistically, quantification of incomplete viral genomes showed that their complementation only partly accounts for the observed reliance on coinfection. Indeed, even when all polymerase components are detected in single-cell mRNA sequencing, robust polymerase activity of GFHK99 virus in mammalian cells is reliant on multiple infection. In sum, IAV collective interactions not only augment reassortment, but can also overcome species-specific barriers to infection. These findings underscore the importance of virus-virus interactions in IAV infection, evolution and emergence.

Introduction

Classically, an infectious unit has been defined as a single virus particle which delivers its genome to a cell, initiates the viral reproductive program, and yields progeny viruses. Increasingly, however, the importance to infection of collective interactions among viruses is being recognized[1-3]. The delivery of multiple virus genomes to a cell allows both antagonistic and mutually beneficial interactions to occur, and these interactions in turn have the potential to shape transmission, pathogenicity, and viral evolutionary pathways.

Recent work has revealed several distinct mechanisms by which multiple viral genomes are co-delivered to a target cell. Diverse taxa including enterovirus, norovirus and rotavirus have all been observed to emerge from cells as groups of particles clustered within extracellular vesicles[4,5]. Adhesion of virus particles to bacterial cell surfaces has a similar clustering effect and increases coinfection of target cells by poliovirus[6]. The aggregation of free virions was found to yield multi-particle infectious units in the case of vesicular stomatitis virus (VSV)[7,8]. Various mechanisms of direct cell-to-cell spread also serve to deliver multiple viral genomes to the same cell[9-11]. The implications of multiple infection in these diverse systems are still being explored. In a number of cases, however, collective delivery was demonstrated to increase the efficiency of infection relative to free virus particles[4,5], or to increase the rate of genetic exchange through recombination[6].

Whether brought about through coordinated infection with physically-linked virions or through independent infection events, the presence of multiple viral genomes within a cell creates the potential for their interaction to alter the course of infection. When distinct variants coinfect, mutually beneficial effects, such as reciprocal compensation for deleterious mutations, can increase overall fitness[12-14]. In the case of IAV, several lines of evidence point to a major role in infection for multiplicity reactivation, the process by which segmented genomes lacking one or more functional segments complement each other[15-19]. Conversely, negative interactions can also arise in which deleteriously mutated genes act in a dominant negative fashion.

Defective interfering particles, which often potently interfere with the production of infectious progeny from a coinfected cell, are the most extreme example of such antagonism[20-22].

Importantly, multiple infection with identical viral genomes can also alter infection outcomes. Such cooperation was documented for VSV and HIV, where rates of transcription and replication were enhanced with increasing multiplicity of infection (MOI)[23,24]. Similarly, faster kinetics of virus production were seen at high MOI for poliovirus and an H3N2 subtype IAV[19,25]. In these instances, it is thought that increased copy number of infecting viral genomes provides a kinetic benefit important in the race to establish infection before innate antiviral responses take hold. Indeed, it has been suggested that multiple infection may be particularly relevant for facilitating viral growth under adverse conditions, such as antiviral drug treatment[3,26].

For IAV, an important adverse condition to consider is that of a novel host environment. IAVs occupy a broad host range, including multiple species of wild waterfowl, poultry, swine, humans and other mammals[27,28]. Host barriers to infection typically confine a given lineage to circulation in one species or a small number of related species[29,30]. Spillovers occur occasionally, however, and can seed novel lineages. When a novel IAV lineage is established in humans, the result is a pandemic of major public health consequence[31,32]. The likelihood of successful cross-species transfer of IAV is determined largely by the presence, absence, and compatibility of host factors on which the virus relies to complete its life cycle, and on the viruses' ability to overcome antiviral defenses in the novel host[33-35].

Owing to the segmented nature of the IAV genome, multiple infection results in viral genetic exchange through reassortment[36,37]. If coinfecting viral genomes are distinct, reassortment will yield genotypic diversity that may facilitate evolution, including adaptation to a new host[38]. Indeed, reassortment involving human seasonal viruses and IAV adapted to non-human hosts was central to the emergence of the last three pandemic strains[39,40]. Thus, among the interactions that occur between coinfecting viruses, reassortment is critical to consider for IAV.

Our objective herein was to assess the degree to which IAV relies on the delivery of multiple viral genomes to a cell to ensure production of progeny. In particular, we sought to determine whether this phenotype varies with host species and with virus strain. We therefore examined multiplicity dependence in avian and mammalian cells for two highly divergent avian-origin IAVs, influenza A/mallard/Minnesota/199106/99 (H3N8) [MaMN99] virus and influenza A/guinea fowl/Hong Kong/WF10/99 (H9N2) [GFHK99] virus. While MaMN99 virus is typical of IAV commonly isolated from wild ducks, GFHK99 virus is representative of the G1 lineage of H9N2 viruses prevalent in the poultry of Southeast Asia, Middle East, and North Africa[41,42]. Results from all virus/cell combinations tested confirm prior reports that cells multiply-infected with IAV produce more viral progeny than singly-infected cells. Importantly, however, the proportion of viral progeny that emerge from coinfected cells varies greatly with virus-host context. The GFHK99 strain exhibits an acute dependence on multiple infection in mammalian cells that is not seen for MaMN99 virus in mammalian cells or for GFHK99 virus in avian cells. The polymerase of the GFHK99 virus drives its host-specific dependence on multiple infection. In line with this finding, both bulk and single-cell measurements of viral RNA showed that polymerase activity in mammalian cells is enhanced with multiple infection. A need for complementation of incomplete viral genomes partially accounts for this cooperative effect. Importantly, however, single cell data indicate that additional multiplicity-dependent mechanisms support RNA synthesis by the GFHK99 virus in mammalian cells. Thus, our data point to an important role for multiple infection in determining the potential for IAV replication in diverse hosts.

Results

Virus-host interactions dictate degree of multiplicity dependence

To evaluate the extent to which IAV relies on multiple infection for productive infection, we initially used coinfection and reassortment as readouts. Reassortment is a useful measure for

coinfection dependence because reassortant viruses must arise from coinfected cells. To ensure accurate quantification of reassortment, coinfections were performed under single-cycle conditions with homologous viruses that differ only by a silent mutation in each segment and the presence of either an HA or HIS epitope tag fused to the HA protein. Such homologous virus pairings were generated in both MaMN99 and GFHK99 strain backgrounds and were named MaMN99 WT / MaMN99 VAR and GFHK99 WT / GFHK99 VAR_1. Tracking of HA and HIS expression by flow cytometry provides a measure of infection that can be compared across cell lines. Quantification of cells expressing one or both epitope tags furthermore gives a means of assessing levels of coinfection across a range of MOIs.

Coinfection and reassortment between homologous viruses of the MaMN99 or GFHK99 strain backgrounds were examined in Madin-Darby canine kidney (MDCK), chicken DF-1 and human A549 cells (**Figure 1, Supplementary Figure 1**). Analysis of MDCK cells infected with the GFHK99 viruses at MOIs ranging from 10 to 0.01 PFU per cell revealed a near linear relationship between total HA^+ cells and dual-HA^+ cells, suggesting that the GFHK99 strain is strictly dependent on multiple infection for HA expression in these cells (**Figure 1A**). HA production resulting from infection with a single strain was more common for GFHK99 in DF-1 cells or MaMN99 in MDCK cells, indicating a lesser dependence on multiple infection (**Figure 1A**). The more dependent expression of *any* HA protein is on coinfection, the more linear the relationship between the percentages of dual-HA^+ cells and HA^+ cells becomes. Conversely, a more quadratic relationship indicates less dependence on coinfection, as individual particles are more often able to express HA independently. We therefore quantified the degree of linearity from the regression models of each dataset and found that only GFHK99 virus in MDCK cells exhibits appreciable linearity in the relationship between HA+ and dual-HA+ cells (**Figure 1B**).

Genotyping of progeny virus from coinfections similarly revealed that reassortment levels vary by virus strain and cell type. In line with observed levels of coinfection, GFHK99 virus exhibits high levels of reassortment in MDCK cells even at low MOIs, indicating that nearly all

progeny virus is produced from WT-VAR_1 coinfected cells (**Figure 1C**). GFHK99 coinfection in A549 cells is also characterized by high levels of reassortment, although less extreme than those seen in MDCK cells (**Supplementary Figure 1**). Compared to GFHK99, MaMN99 viruses infecting MDCK cells show lower levels of reassortment (**Figure 1C**). Moreover, reassortment of GFHK99 viruses is markedly reduced in DF-1 cells compared to that seen in MDCK cells (**Figure 1D**). These results clearly reveal differing degrees of multiplicity dependence for different virus/cell pairings and therefore indicate that multiple infection dependence, rather than solely being an intrinsic property of a virus strain, is determined through virus-host interactions.

That all virus/cell pairings tested show evidence of multiplicity dependence is highlighted by comparison of the experimental reassortment data to a theoretical prediction that assumes infection is perfectly efficient (**Figure 1C, 1D**). This theoretical prediction was published previously[18] and is derived from a computational model in which the number of viral progeny produced by an infected cell is constant. Because singly- and multiply-infected cells make equivalent numbers progeny, reassortment is predicted to increase only gradually at low levels of infection (low %HA+) where coinfection is relatively rare. By contrast, reassortment observed experimentally reaches high levels much more rapidly. High reassortment indicates that viral progeny production is focused in the proportion of the infected cell population that is multiply infected. Coinfection dependence, therefore, is evident in all virus-cell pairings, but particularly strong for GFHK99 in MDCK cells.

Strain and host specific phenotypes are also evident **in vivo**

To determine whether host-dependent reliance on multiple infection extended to *in vivo* infection, we performed coinfections with MaMN99 WT and VAR viruses in guinea pigs and GFHK99 WT and VAR_1 viruses in guinea pigs and quail. To ensure use of comparable effective doses for each virus/host pairing, the 50% infectious dose (ID_{50}) of each virus mixture was first determined experimentally in the animal models used. Guinea pigs were then infected

intranasally with 10^2 $GPID_{50}$ of either the GFHK99 or MaMN99 WT/VAR mixture and nasal washes were collected daily. Japanese quail were infected with 10^2 QID_{50} of the GFHK99 virus mixture via an oculo-naso-tracheal route and tracheal swabs were collected daily. To evaluate the frequency of reassortment, plaque isolates from these upper respiratory samples were genotyped for each animal on each day. Because multicycle replication *in vivo* allows the propagation of reassortants, analysis of genotypic diversity rather than percent reassortment is more informative for these experiments. Thus, the effective diversity (Hill's N_2) was calculated for each dataset and plotted as a function of time post-inoculation (**Figure 1E, 1F**). The viruses collected from GFHK99 infected guinea pigs show much higher genotypic diversity throughout the course of infection than viruses isolated from MaMN99 infected guinea pigs (**Figure 1E**) or GFHK99 infected quail (**Figure 1F**). These data indicate that the virus-host interactions which determine dependence on multiple infection in cell culture extend to *in vivo* infection.

Multiple infection enhances viral growth

The abundant reassortment observed with GFHK99 viruses in mammalian systems suggests that multiple infection plays a major role in determining the productivity of an infected cell. We therefore hypothesized that increasing MOI would augment the burst size, or viral output, of infected cells and that the magnitude of this effect would be greater for GFHK99 in MDCK cells than for GFHK99 in DF-1 cells or MaMN99 in MDCK cells. To test this prediction, we infected over a range of MOIs with the same mixtures of GFHK99 WT and VAR_1 or MaMN99 WT and VAR viruses used above and then measured PFU produced per cell under single-cycle conditions. Under non-saturating conditions (determined by flow cytometry to be MOI <1 PFU per cell, as shown in **Supplementary Figure 2**), increasing MOI resulted in accelerated viral growth and higher burst size for all three virus/cell pairings (**Figure 2A-C**). As predicted, however, increasing the MOI of GFHK99 in MDCK cells resulted in a further enhancement of

193 viral amplification (**Figure 2D**). Thus, the benefit conferred by multiple infection was greater for
194 GFHK99 in MDCK cells compared to either MaMN99 in MDCK cells or GFHK99 in DF-1 cells.
195 We reasoned that the cooperative effect observed might result from i) complementation
196 of incomplete viral genomes or ii) a benefit of increased viral genome copy number per cell. In
197 an effort to differentiate between these possibilities, we measured growth of GFHK99 in MDCK
198 and DF-1 cells infected at a range of MOIs greater than 1 PFU per cell. Because these
199 conditions are saturating (**Supplementary Figure 2**), incomplete viral genomes are unlikely to
200 be prevalent and any benefit of increasing MOI would be attributable to increasing genome copy
201 numbers per cell. In both cell types, MOIs between 1 and 20 PFU per cell result in similar peak
202 viral titers (**Figure 2E-F**). This saturation of cooperation at higher MOIs suggests diminishing
203 returns from additional genome copies above a certain threshold. Calculation of fold change in
204 viral amplification revealed that burst size was either unchanged or negatively affected by
205 increasing MOI above 1 PFU per cell (**Figure 2G**). Whether this threshold is imposed by a need
206 for complementation or another mechanism sensitive to saturation remained unclear. Overall,
207 however, the increase in viral amplification with increased multiplicity at sub-saturating MOIs
208 strengthened our prior conclusion that viral growth, and particularly productivity of GFHK99 virus
209 in MDCK cells, is enhanced by multiple infection.
210
211 ***The viral polymerase is a major determinant of multiple infection dependence***
212 To identify viral genetic determinants of multiple infection dependence, we mapped segments
213 responsible for the high reassortment phenotype of GFHK99 in MDCK cells. Reverse genetics
214 was used to place one or more genes from GFHK99 into a MaMN99 background. We created a
215 panel of chimeras containing the HA, NP, or the full polymerase complex and NP (3PNP) of
216 GFHK99 virus in the MaMN99 background. We also generated the reciprocal swap of this last
217 genotype in which NS, M, NA and HA segments were derived from GFHK99. These segment
218 groupings were selected for exchange based on their functions in the viral life cycle. For each

chimeric genotype, homologous WT and VAR strains were generated to allow tracking of homologous reassortment.

Coinfections with matched WT and VAR strains were performed in MDCK cells and HA expression and reassortment were measured as in **Figure 1**. When levels of dual HA positivity are assessed, most MaMN99:GFHK99 chimeric genotypes cluster together with the parental MaMN99 virus, suggesting a relatively low dependence on multiple infection for HA expression (**Figure 3A-B**). By contrast, the MaMN99:GFHK99 3PNP genotype gives results similar to those of the parental GFHK99 (**Figure 3A-B**). Quantification of reassortment revealed that all chimeric viruses reassort at a higher frequency than MaMN99 parental strains but that the MaMN99:GFHK99 3PNP viruses show the highest reassortment, comparable to that seen for the parental GFHK99 genotype (**Figure 3C-D**). Thus, while other viral genes may make minor contributions, the viral polymerase is the primary genetic determinant of the high reassortment exhibited by GFHK99 in MDCK cells and defines a need for cooperation between coinfecting viruses.

Multiple infection enhances viral RNA replication

Because genetic mapping of the GFHK99 high reassortment phenotype implicated the viral polymerase, we sought to ascertain the effects of multiple infection on polymerase function. We therefore measured GFHK99 WT viral RNA synthesis in the absence and presence of increasing amounts of a homologous coinfecting virus. To evaluate host specificity, we did this analysis in both MDCK and DF-1 cells. The coinfecting virus, GFHK99 VAR_2, was generated in the GFHK99 background to avoid genetic incompatibility and carries silent mutations in each segment that disrupt primer binding sites. Cells were infected with low MOI (0.005 PFU per cell) of GFHK99 WT virus to ensure receipt of a single copy of the virus genome. Concurrently, cells were infected with increasing doses of GFHK99 VAR_2 virus. Digital droplet PCR (ddPCR) with primers specific for GFHK99 WT cDNA was then used to quantify replication of WT genomes.

The results show that, in both cell types, coinfection with low to moderate doses of the VAR_2 virus increases levels of GFHK99 WT vRNA (**Figure 4A**). At the highest doses of VAR_2 virus used, however, a suppressive effect is observed. Importantly, the amount of coinfection required to reach maximal vRNA production differs among cell lines: in MDCK cells 10-fold more VAR_2 virus (1 PFU per cell) is needed than in DF-1 cells (0.1 PFU per cell). The maximal impact of VAR_2 virus on WT vRNA production is also greater in MDCK cells: a ~60-fold enhancement is seen, compared to only ~2-fold in DF-1 cells.

To verify these observations in a more physiologically relevant system, we repeated the experiment in primary human tracheobronchial epithelial (HTBE) cells differentiated at an air-liquid interface. Similar to MDCK cells, these primary human cells exhibit maximal GFHK99 WT RNA production with addition of 1 PFU per cell GFHK99 VAR_2 virus (**Figure 4B**). Peak RNA replication is >10 fold higher in HTBE cells than without coinfecting virus.

Thus, in all three cell types tested, the introduction of coinfecting VAR_2 virus reveals a cooperative effect acting at the level of RNA synthesis. At very high doses of VAR_2 virus, WT RNA levels decline, suggesting competition for a limited resource at these extreme MOIs. Most notably, the magnitude of the cooperative effect and the amount of VAR_2 virus needed to reach maximal WT RNA levels are much greater in mammalian cells than avian cells. These differing outcomes indicate that the multiplicity dependence of GFHK99 polymerase function is modulated by host factors that differ between mammalian and avian hosts.

Multiple infection accelerates viral replication and transcription

To more finely assess the effects of multiplicity on polymerase function in various virus-host combinations, we measured vRNA, mRNA, and cRNA over time following low or high MOI infection (**Figure 5A**). MaMN99 and GFHK99 viruses were examined in MDCK cells and GFHK99 virus in DF-1 cells. To evaluate the activity of the viral polymerase when the encoding genes are supplied as low or single copies, a dose of 0.5 RNA copies per cell was used for

infection. Under these low MOI conditions, all three viral RNA species accumulate at a significantly higher rate for GFHK99 in DF-1 cells and MaMN99 in MDCK cells than for GFHK99 in MDCK cells (**Figure 5B**). In defining a high MOI dose, we elected to use HA expressing units, as determined by flow cytometry, rather than genome copy number (**Supplementary Figure 3**). This measure gives a functional readout for polymerase activity, and therefore allows a dose to be chosen that ensures the vast majority of cells carry an active viral polymerase. The high MOI dose used was 3.0 HA expressing units per cell. At this high MOI, accumulation of GFHK99 mRNA, vRNA and cRNA in MDCK cells occurs at a similar rate to that seen for GFHK99 in DF-1 cells or MaMN99 in MDCK cells (**Figure 5C**). Thus, a host-specific defect in GFHK99 polymerase activity that affects synthesis of all three viral RNA species is seen at low MOI. This defect is, however, resolved under conditions where multiple infection is prevalent.

Single cell mRNA sequencing reveals a need for cooperation beyond complementation

An important limitation of working with bulk RNA extracted from a population of cells is the inability to distinguish between i) low, but uniform, RNA synthesis in all cells and ii) robust RNA synthesis in only a minority of cells. To elucidate the basis for cooperation at higher MOIs, it was important to determine which of these scenarios gives rise to the low average viral RNA levels that characterize low MOI infection with GFHK99 in MDCK cells. A highly heterogeneous picture, with abundant viral RNAs in a minority of cells, would be expected if incomplete viral genomes are common but cells with a complete set of polymerase genes support robust viral RNA synthesis. Conversely, uniformly low levels of viral products would be expected if even complete viral genomes cannot support robust polymerase activity in the context of singular infection.

To evaluate the heterogeneity of viral RNA synthesis at the single cell level, we used single-cell mRNA sequencing. We infected DF-1 or MDCK cells with GFHK99 virus under single-cycle conditions and collected cells at 8 h post-infection for mRNA barcoding on the 10X

297 Genomics Chromium platform prior to sequencing. The relative abundance of mRNA from each
298 viral transcript was calculated by normalizing to the median number of transcripts per cell in that
299 infection. Cells in which at least one viral mRNA molecule was detected were analyzed further.
300 The number of cells that met this criterion ranged between 182 and 478 per infection condition
301 (MOI and cell type combination). We found that the amount of detected GFHK99 viral mRNA
302 varies widely between individual DF-1 cells (**Figure 6A**), which is consistent with previous
303 observations[43]. In contrast, GFHK99 viral mRNA levels are uniformly low in MDCK cells under
304 the relatively low MOI conditions used.

305 Because only subset of a cell's transcripts is captured and therefore reliably detected[44],
306 the 10X platform does not allow a robust determination of segment presence or absence in a
307 cell. Where viral mRNAs derived from a given segment are detected, however, one can
308 conclude that the corresponding vRNA was present. To evaluate whether low transcript
309 abundance corresponded to the lack of one or more polymerase-encoding segments, we
310 therefore stratified the data based on detection of all four segments necessary to support
311 transcription (PB2, PB1, PA and NP) (**Figure 6A**). Viral transcript levels are markedly increased
312 in DF-1 cells that contained the PB2, PB1, PA, and NP segments compared to those in which
313 one or more of these segments was not detected. Averaging across all MOIs, a 10-fold increase
314 in transcript abundance was noted in DF-1 cells ($p < 10^{-16}$, linear mixed effects model). In
315 contrast, viral transcription in MDCK cells is consistently low, and the presence of polymerase
316 complex confers no benefit (**Figure 6A**).

317 Data presented above from bulk samples indicate that multiple infection is needed for
318 efficient GFHK99 transcription in MDCK cells. To measure the impact of coinfecting virus in
319 individual cells, we repeated the single-cell sequencing experiment with the addition of
320 genetically marked variants of GFHK99 virus. For an mRNA sequencing assay, marker
321 mutations proximal to the poly-A tail of the viral transcripts are needed; we therefore generated
322 variant viruses that carry synonymous nucleotide changes near the 5' end of each vRNA,

GFHK99 $mVAR_1$ and GFHK99 $mVAR_2$. Cells were inoculated with GFHK99 WT and GFHK99 $mVAR_1$ viruses in a 1:1 ratio and the combined MOI was the same as that used for GFHK99 WT in the first experiment. Coinfection with GFHK99 $mVAR_2$ virus was performed simultaneously and this virus was used at the concentration found to be optimal for WT viral RNA replication in **Figure 4A**; the MOI therefore differed between DF-1 (0.1 PFU per cell) and MDCK (1.0 PFU per cell) cells. After mRNA sequencing, cells in which transcripts from all eight $mVAR_2$ virus segments were detected were analyzed further. Between 131 and 240 cells per infection condition (MOI, cell type, virus strain) met this criterion. The viral transcript levels per cell detected in this second experiment are shown in **Figure 6B** alongside data from the first experiment for comparison. In this figure, GFHK99 WT and GFHK99 $mVAR_1$ mRNAs are plotted separately; the concordance between these two datasets gives an indicator of reproducibility. In comparing the two infections, we observe that total viral transcript abundance is 72% lower in MDCK cells compared to DF-1 cells in the first infection ($p < 10^{-16}$, linear mixed effects model), but this effect is almost entirely mitigated by the presence of $mVAR_2$ virus, as transcript abundance is reduced by only 11% in MDCK cells in the second infection ($p < 10^{-16}$, linear mixed effects model). This reduction in the disparity between DF-1 and MDCK cell viral transcript abundance resulted from the fact that $mVAR_2$ virus increased transcript abundance by 102% in DF-1 cells, but 545% in MDCK cells ($p < 10^{-16}$, linear mixed effects model) (**Figure 6B)**. These data underscore the significance of viral collective interaction to ensure productive infection in diverse hosts.

Frequency of incomplete GFHK99 genomes in MDCK cells is moderate

Our single-cell sequencing results suggest that the presence of a complete viral genome in the infected cell is not sufficient to support robust transcription of GFHK99 vRNAs. All eight viral gene segments are, however, necessary for productive infection[45] and could play an important role in the reliance of GFHK99 virus on multiple infection in mammalian systems. We therefore

sought to quantify the frequency with which fewer than eight vRNAs are replicated in GFHK99 infected MDCK cells. Given the sensitivity limitations of the single-cell mRNA sequencing method, we employed a single-cell assay that we designed previously for this purpose[19]. MDCK cells were coinfected with a low MOI of GFHK99 WT virus and a high MOI of GFHK99 VAR_2 virus. The GFHK99 VAR_2 virus acts to ensure propagation of the WT virus gene segments, even when less than the full WT viral genome is available for transcription and replication. Following inoculation, single cells were sorted into wells which contain a naïve cell monolayer and multicycle replication was allowed for 48 h. To determine which viral gene segments were present in the initially sorted cell, RT-qPCR with primers that differentiate WT and VAR_2 gene segments was applied. As detailed in the Methods, the frequencies of VAR_2 virus infection, WT virus infection, and each distinct WT segment were used to estimate the probability that a cell infected with a single WT virus would contain a given segment. We termed the resultant parameter Probability Present (P_P). The experimentally determined P_P values vary among the segments, with a range of 0.57 to 0.88 (**Figure 7A**). The product of the eight P_P values gives an estimate of the proportion of singular infections in which all eight segments are available for replication. This estimate is 6.5% for GFHK99 in MDCK cells.

The high reassortment of GFHK99 WT and VAR_1 viruses in MDCK cells indicates that progeny viruses predominantly originate from multiply infected cells in this system. To evaluate whether incomplete viral genomes account for this focusing of GFHK99 virus production within multiply infected cells, we used our previously published computational model of IAV coinfection and reassortment[18]. In this model, the frequency of segment delivery upon replication is governed by eight P_P parameters and an infected cell only produces virus if at least one copy of all eight segments are present. Importantly, in this model the amount of virus produced from productively infected cells is constant – there is no additional benefit to multiple infection. When the eight experimentally determined P_P values for GFHK99 virus in MDCK cells are used to parameterize the model, the theoretical prediction of reassortment frequency is much lower than

375 that observed experimentally for GFHK99 WT and VAR_1 viruses in MDCK cells (**Figure 7B**).
376 This discrepancy indicates that the frequency of missing segments cannot fully account for the
377 high reassortment seen. Thus, the collective interactions on which the GFHK99 virus relies for
378 replication in mammalian systems appears to extend beyond complementation. A full viral
379 genome is necessary, but not sufficient, to support robust replication.
380
381 **Discussion**
382 Using small genetic tags and a range of molecular tools for their detection, we investigated the
383 determinants of and mechanistic basis for IAV multiplicity dependence. Our data reveal that
384 both viral and host features dictate the degree to which productive IAV infection relies on
385 cooperation. Thus, multiple infection dependence is a property determined through interaction
386 between the virus and the infected cell, rather than an intrinsic property of the virus. Differences
387 between virus strains and host systems in multiple infection dependence lead to phenotypic
388 differences in the amount of reassortment that occurs upon coinfection. The demonstration of
389 these reassortment differences in mammalian and avian models points to the relevance of viral
390 collective interactions for IAV evolution and emergence. Mechanistically, our data indicate that
391 multiple infection is needed in part for complementation of incomplete viral genomes, but that
392 such complementation is not sufficient to ensure productive infection in all virus-host systems.
393 Rather, we see that the GFHK99 polymerase requires a second form of cooperation to support
394 efficient RNA synthesis in mammalian cells. For this reason, robust infection of GFHK99 virus
395 in mammalian systems is achieved only in the context of high MOI infection. Thus, the data
396 presented reveal that infection efficiency and the need for cooperation varies with virus-host
397 context, and the viral polymerase is a major driver of this phenotype.
398 An important implication of viral genome segmentation is the potential for replication of
399 incomplete genomes within infected cells[2]. For most segmented viruses of vertebrates, including
400 IAV, each segment encodes at least one essential gene product and a genome lacking one or

more functional segments cannot support the production of progeny viruses. Complementation is therefore a major class of collective interaction for viruses with segmented genomes, the relevance of which likely depends on the extent to which delivery and replication of the various genome segments is coordinated for a given virus species[46]. For IAV, we and others have demonstrated that, within singly-infected cells, a subset of segments fails to be replicated or expressed with high frequency[16,19,43]. Specifically, for influenza A/Panama/2007/99 (H3N2) virus in MDCK cells, we found that delivery of a single viral genome results in replication of all eight segments only 1.2% of the time[19]. Data reported herein for GFHK99 virus indicate that a somewhat higher proportion of replicated viral genomes are complete – namely, 6.5%. Thus, GFHK99 virus is partially dependent on complementation for productive infection in this cell line. However, the high levels of reassortment seen between GFHK99 WT and VAR_1 viruses in mammalian cells indicate that additional cooperative interactions are at play. This is made clear by the discrepancy between observed GFHK99 virus reassortment and the reassortment levels expected if complementation is the only cooperative effect considered. A necessary but insufficient role for complete viral genomes is further supported by the results of single cell mRNA sequencing of GFHK99 virus infected MDCK cells. Here, viral transcripts are produced at low copy numbers even when all four segments needed to support viral RNA synthesis are confirmed to be present. This outcome is in contrast to that observed for GFHK99 virus in DF-1 cells and to that reported previously for WSN virus in A549 cells, where the heterogeneity in viral transcript levels among cells could be attributed in part to the apparent absence of one or more polymerase complex genes[43]. Notably, however, the restriction of GFHK99 viral transcription in MDCK cells was largely mitigated by the addition of a homologous coinfecting virus. These data point to a model in which the presence of not just complete genomes, but rather multiple copies of the viral genome, are needed to overcome host-specific barriers to GFHK99 infection in mammalian systems.

Insight into the nature of this second cooperative interaction is gleaned from the observation that the amount of viral RNA produced from a constant input of GFHK99 WT viral genomes is significantly increased with the addition of a homologous virus that is genetically tagged to allow independent detection. Because GFHK99 WT virus RNAs can be quantified separately from the coinfecting VAR_2 virus RNAs, we can conclude that the coinfecting virus functions in *trans* to support GFHK99 WT virus replication. This interaction is likely to occur at the protein level, with increased genome copy number supporting the expression of higher levels of viral polymerase proteins or cofactors. This proposed mechanism is supported by prior work showing that the IAV polymerase can act in *trans* to propagate temperature sensitive (ts) variants at non-permissive temperatures[47].

Our data implicate the viral polymerase in defining an acute reliance on cooperation for efficient viral RNA synthesis and viral progeny production. While our experiments focused on only two avian IAVs, it is well known that avian-adapted IAV polymerases require adaptive changes for efficient replication in mammalian cells[33,48,49]. The conformation or composition of the GFHK99 viral polymerase may lead to defects in transcription or replication due to poor interactions with mammalian host factors, such as ANP32A[50]. Low functionality of the viral polymerase complex may furthermore lead to the synthesis of abortive products, such as mini viral RNAs[51]. Thus, the multiplicity dependence of GFHK99 in mammalian systems may be a manifestation of poor adaptation of the viral polymerase to the host cell. Importantly, however, it appears that this lack of adaptation can be at least partially overcome when multiple viral genomes are delivered to the same cell. While MaMN99 virus is also not adapted to mammalian systems, it is only distantly related to GFHK99 virus and is representative of viruses that circulate in a taxonomically and geographically distinct avian population compared to the poultry hosts of GFHK99 virus. It will be important in future studies to delineate further the IAV lineages and host contexts in which an acute need for cooperation exists.

The clear involvement of the polymerase does not exclude the possibility that other virus-host interactions may impact the need for cooperation. In fact, reassortment levels measured for chimeric GFHK99-MaMN99 viruses indicated that other viral components contribute to the high reassortment phenotype of GFHK99 virus. For example, it has been postulated that the pH of fusion of HA, which dictates when the viral genome is released from endosomes, determines the amount of time that viral gene segments are vulnerable to diffusion or degradation during transit to the nucleus [52]. Because stochastic loss of a subset of gene segments prior to nuclear import would likely be overcome through multiple infection, HA pH of fusion may determine the need for cooperation in some virus-host contexts. The contribution of particular viral proteins to coinfection dependence is relevant for understanding barriers to zoonotic infection and predicting the likelihood of reassortment following zoonoses.

The H9N2 subtype is of particular relevance in the context of zoonotic infection as viruses of this subtype are highly prevalent at the poultry-human interface, sporadic human infections have been reported, and H9N2 viruses share several related genes with H5N1 and H7N9 subtype viruses that have caused hundreds of severe human infections[53-58]. The G1 lineage to which the GFHK99 virus belongs circulates widely in the poultry of Southeast Asia and North Africa and reassorts frequently with other poultry adapted IAVs[41,42,59,60]. The prevalence of reassortment suggests that the internal gene segments – which comprise the six non-HA, non-NA segments – are compatible with other genotypes. Our comparison of reassortment in guinea pigs and quail furthermore indicates that reassortment could be particularly prevalent in the context of zoonotic infection of mammals. This phenotype of high reassortment in mammals is expected to extend to the H5N1 and H7N9 subtype viruses of public health concern, which carry polymerase genes related to those of the GFHK99 virus[56-58]. While reassortment is typically deleterious owing to negative epistasis among heterologous segments[61,62], a high frequency of reassortment creates greater opportunity for fit genotypes to

476 arise and adapt, and should therefore be considered in assessing the risk of emergence posed
477 by non-human adapted IAVs.

478 Our work reveals an underappreciated facet of virus-host interactions: the extent to
479 which IAV relies on cooperation with coinfecting viruses is both strain and host dependent.
480 Varied phenotypes of multiplicity dependence occurring in different virus-host contexts likely
481 have important implications for viral fitness and viral evolution. Differences in coinfection
482 dependence are expected to lead to differences in the viral dose required to establish a new
483 infection, which in turn has implications for both the likelihood of transmission and the
484 predominant mode of transmission. For example, transmission among close contacts is
485 associated with the transfer of higher viral loads[63]. Reliance on cooperation is also expected to
486 impact the spatial dynamics of viral spread within an individual[19]. For example, long-distance
487 dispersal of virus within a host is less likely to be productive in a system where the virus is
488 highly dependent on cooperative interactions[64]. Finally, the features which impact coinfection
489 dependence are also likely to impact viral evolution by changing how a virus population samples
490 the available sequence space. As discussed above, multiplicity dependence increases the
491 opportunity for genetic exchange through reassortment, which may in turn slow the
492 accumulation of deleterious mutations and allow coupling of advantageous mutations[65]. A need
493 for cooperation would also be predicted to increase the likelihood that less fit variants are
494 propagated as a result of phenotypic hiding in coinfected cells[66-68]. Thus, host and strain
495 specificity in multiple infection dependence are likely to play an important role in determining the
496 outcomes of IAV infection and evolution in diverse hosts.

497

498 **Acknowledgements**

499 We thank David Stallknecht (University of Georgia) for providing the A/mallard/MN/199106/99
500 (H3N8) biological isolate. We thank Hui Tao, Shamika Danzy, and Ginger Geiger for technical
501 assistance. This work was funded in part by NIH R01 AI127799 (to ACL and DRP); NIH/NIAID

Centers of Excellence in Influenza Research and Surveillance (CEIRS), contract numbers HHSN272201400004C (to ACL) and HHSN272201400008C (to DRP). Additional funds were provided by the Georgia Research Alliance and the Georgia Poultry Federation (to DRP) and NIH/NIAID Genomic Centers for Infectious Diseases (GCID), Award Number U19AI110819.

Methods

Cells

Madin-Darby canine kidney (MDCK) cells, a gift from Peter Palese, Icahn School of Medicine at Mount Sinai were used in coinfection experiments, growth curves, and dosage experiments with increasing amounts of GFHK99 VAR_2 virus added. MDCK cells from Daniel Perez at University of Georgia were used for plaque assays as this variant of the MDCK line was found to yield more distinct plaques for the GFHK99 strain. Both MDCK cell lines were maintained in minimal essential medium (MEM; Gibco) supplemented with 10% fetal bovine serum (FBS; Atlanta Biologicals), penicillin (100 IU), and streptomycin (100 µg per mL) (PS; Corning). A549 cells (ATCC CCL-185) were maintained in F-12K nutrient mixture with L-glutamine (Corning) supplemented with 10% FBS and PS. 293T cells (ATCC CRL-3216) and DF-1 cells (ATCC CRL-12203) were maintained in Dulbecco's minimal essential medium (DMEM; Gibco) supplemented with 10% FBS and PS. Human tracheobronchial epithelial (HTBE) cells from a single donor were acquired from Lonza and were amplified and differentiated into air-liquid interface cultures as recommended by Lonza and described by Danzy et al.[69]. All cells were cultured at 37°C and 5% CO_2 in a humidified incubator.

Viruses

All viruses used in this study were generated through reverse genetics[70]. 293T cells transfected with reverse genetics plasmids 16-24 h prior were injected into the allantoic cavity of 9-11 day old embryonated chicken eggs and incubated at 37°C for 40-48 h. The resultant egg passage 1

528 stocks were used in experiments. Defective interfering segment content of PB2, PB1, and PA
529 segments was confirmed to be minimal for each virus stock, following a method described
530 previously[71]. The reverse genetics system for influenza A/guinea fowl/Hong Kong/WF10/99
531 (H9N2) virus was reported previously[72,73]. This strain has been referred to as WF10 in previous
532 publications[72-74] but, for consistency with other strains used in the present manuscript, is
533 referred to herein as GFHK99. A low passage isolate of influenza
534 A/mallard/Minnesota/199106/99 (H3N8) virus, referred to herein as MaMN99, was obtained
535 from David Stallknecht at the University of Georgia[75]. The virus was passaged once eggs and
536 then the eight cDNAs were generated and cloned into the pDP2002 vector[76]. To increase the
537 efficiency of virus recovery for rescues containing polymerase components from the MaMN99
538 virus, pCAGGS support plasmids encoding PB2, PB1, PA, and NP proteins of the A/WSN/33
539 (H1N1) strain were supplied.

540 GFHK99 WT and MaMN99 VAR viruses were engineered to contain a 6XHis epitope tag
541 plus GGGS linker at the N terminus of the HA protein following the signal peptide. GFHK99
542 VAR_1 and MaMN99 WT viruses contain similarly modified HA genes, with an HA epitope tag
543 plus a GGGS linker inserted at the N terminus of the HA protein[77].

544 Silent mutations introduced by site directed mutagenesis were used to confer altered
545 melting properties to allow high resolution melt genotyping of WT and VAR_1 segment origin.
546 Mutations introduced in to GFHK99 VAR_1 and MaMN99 VAR viruses are listed in
547 **Supplementary Table 1**. Mutations introduced into the GFHK99 VAR_2 strain were designed to
548 confer unique primer binding sites relative to GFHK99 WT virus for use in digital PCR-based
549 genotyping. The mutations introduced are also listed in **Supplementary Table 1**. Viruses used
550 for single cell mRNA sequencing, GFHK99 $mVAR_1$ and GFHK99 $mVAR_2$, were generated from
551 the GFHK99 WT strain with no HIS tag, following the approach described in Russell et al.[43]. The
552 mutations introduced were designed to be detected where sequence data is available from only
553 the 3' end of each transcript. Site directed mutagenesis was therefore used to place two silent

554 mutations proximal to the stop codon in each viral cDNA. The mutations allow differentiation
555 among the segments of the three strains. All such mutations are reported in **Supplementary**
556 **Table 1**.
557
558 ***Coinfection in cultured cells for quantification of coinfection and reassortment***
559 MDCK, DF-1, or A549 cells were seeded at a density of 4 x10^5 cells per well in 6-well dishes 24
560 h before inoculation. Virus inoculum was prepared by combining WT and VAR viruses at high
561 titer in a 1:1 ratio based on PFU titers, and then diluting in PBS to achieve MOIs ranging from
562 10 to 0.01 PFU per cell. Synchronized infection conditions were used, as follows. Cell
563 monolayers were washed three times with PBS and placed on ice. Chilled virus inoculum was
564 added to each well at a 250 µL volume and incubated at 4°C for 45 minutes with occasional
565 rocking. Inoculum was aspirated and cell monolayer was rinsed three times with cold PBS
566 before addition of warm virus medium. Due to low viral growth of GFHK99 virus in DF-1 cells,
567 acid inactivation of inoculum virus was performed at 1 h post-infection for this cell type. For acid
568 inactivation, media was aspirated and replaced with 500 µL of PBS-HCl, pH 3.00 and incubated
569 5 min at 37°C. Cells were then washed once with PBS before the addition of virus medium. At 3
570 h post-infection, virus medium was replaced with ammonium chloride-containing virus medium.
571 GFHK99 virus infected cells were harvested at 12 h post infection due to high amounts of CPE
572 at later time points. Cells infected with MaMN99 virus and MaMN99:GFHK99 chimeric viruses
573 were harvested at 16 h post-infection. Virus medium for each cell line was prepared by
574 supplementing the appropriate media (MDCK, MEM; DF1, DMEM; A549, F12K) with 4.3%
575 bovine serum albumin and penicillin (100 IU), and streptomycin (100 µg per mL). Ammonium
576 chloride-containing virus medium was prepared by the addition of HEPES buffer and NH_4Cl at
577 final concentrations of 50 mM and 20 mM, respectively, to virus media.
578

579 ***Determination of infection levels based on HA surface expression.*** To enumerate infected
580 cells, surface expression of HIS and HA epitope tags was detected by flow cytometry. This
581 method was previously described in detail[77]. The percentage of cells that were positive for either
582 or both epitope tags is expressed as percentage of cells HA^+. The percentage of cells that were
583 positive for both epitope tags is expressed as percentage of cells dual-HA^+. The relationship
584 between these two parameters was evaluated by plotting % cells dual-HA^+ against % cells HA^+
585 and regressing the resultant curve as a quadratic polynomial (*% cells dual-HA*$^+$ = β_2*(*% cells*
586 *HA*$^+$)^2 + β_1*(*% cells HA*$^+$), where β_2 and β_1 are genotype-specific). From the regression models,
587 we then quantified the degree of linearity using the equation $\%\ linearity = \frac{|\beta_1|}{|\beta_1|+|\beta_2|}$.

588

589 ***Animal models and reassortment*** **in vivo**

590 Quail eggs obtained from the College of Veterinary Medicine, University of Georgia, were
591 hatched at the Poultry Diagnostic and Research Center, University of Georgia. Two days before
592 virus inoculation, quail sera were confirmed to be seronegative for IAV exposure by NP ELISA
593 (IDEXX, Westbrook, ME). At 3-weeks of age, birds were moved into a HEPA in/out BSL2 facility
594 and each group divided into individual isolator units.

595 Groups (n=6) of 3-week old Japanese quail (Coturnix Japonica) were used to determine
596 the 50% quail infectious dose of the 1:1 GFHK99 WT and GFHK99 VAR_1 virus mixture. Each
597 quail was inoculated with 500 µl by oculo-naso-tracheal route of virus mixture in PBS, at
598 increasing concentrations of 10^0 to 10^6 $TCID_{50}$ per 500 µL. Tracheal and cloacal swab
599 specimens were collected daily from each bird in brain heart infusion media (BHI). Swab
600 samples were analyzed by $TCID_{50}$ assay and titers of tracheal swabs collected at 4 d post-
601 inoculation were used to determine the QID_{50} by the Reed and Muench method[78]. Virus was not
602 detected in cloacal swabs. QID_{50} was found to be equivalent to 1 $TCID_{50}$.

603 To quantify reassortment in quail, samples collected from quail (n=6) infected with the
604 10^2 $TCID_{50}$ dose of the 1:1 GFHK99 WT and GFHK99 VAR_1 virus mixture were used. These
605 were the same birds as used to determine QID_{50}. Virus shedding kinetics were determined by
606 plaque assay of tracheal swab samples and samples from days 1, 3, and 5 were chosen for
607 genotyping of virus isolates.

608 Female Hartley strain guinea pigs weighing 250-350 g were obtained from Charles River
609 Laboratories. The $GPID_{50}$ of GFHK99 WT/VAR_1 and MaMN99 WT/VAR virus mixtures were
610 determined as follows. Groups of four guinea pigs were inoculated intranasally with virus
611 mixture in PBS at doses of 10^0 to 10^5 PFU per 300 µL inoculum. Daily nasal washes were
612 collected in 1 mL PBS and titered by plaque assay. Results from day 2 nasal washes were used
613 to determine the $GPID_{50}$ by the Reed and Muench method[78]. The $GPID_{50}$ of GFHK99 virus was
614 found to be 2.1 x 10^3 PFU, while that of MaMN99 virus was determined to be 2.1 x 10^1 PFU.

615 To evaluate reassortment kinetics in guinea pigs, groups of six animals were infected
616 with 10^2 x $GPID_{50}$ of the aforementioned GFHK99 WT / VAR_1 virus mixture or the MaMN99 WT /
617 VAR virus mixture. Virus inoculum was given intrasnasally in a 300 µl volume of PBS. Nasal
618 washes were performed on days 1-6 post-inoculation and titered for viral shedding by plaque
619 assay. HRM genotyping was performed on samples collected on day 1, 3, and 5 for each
620 guinea pig.

621

622 ***Quantification of reassortment and effective diversity***

623 Reassortment was quantified for *in vitro* coinfection supernatants, guinea pig nasal washes, and
624 quail tracheal swabs as described previously[77]. Briefly, plaque assays were performed in 10 cm
625 dishes to isolate virus clones. 1 mL serological pipettes were used to collect agar plugs into 160
626 µl PBS. Using a ZR-96 viral RNA kit (Zymo), RNA was extracted from the agar plugs and eluted
627 in 40 µl nuclease free water (Invitrogen). Reverse transcription was performed using Maxima RT
628 (Thermofisher) according to the manufacturer's protocol. The resulting cDNA was diluted 1:4 in

nuclease free water and each cDNA was combined with segment specific primers and Precision Melt Supermix (Bio-Rad) and analyzed by qPCR in a CFX384 Touch real-time PCR detection system (Bio-Rad) designed to amplify a ~100 bp region of each gene segment which contains a single nucleotide change in the VAR virus. The qPCR was followed by high-resolution melt (HRM) analysis to differentiate WT and VAR amplicons[79]. Precision Melt Analysis software (Bio-Rad) was used to determine the parental virus origin of each gene segment based on melting properties of the cDNAs and comparison to WT and VAR controls.

Viral genotypic diversity was quantified as reported previously[80] by calculating Simpson's Index, given by $D = \text{sum}(p_i^2)$, where p_i represents the proportional abundance of each genotype[81]. Simpson's Index accounts for both the raw number of species and variation in abundance of each, and is sensitive to the abundance of dominant species. Because Simpson's Index does not scale linearly, each sample's Simpson's Index value was converted to a corresponding Hill number to derive its effective diversity, $N_2 = 1/D$[82], which is defined as the number of equally abundant species required to generate the observed diversity in a sample community. Because it scales linearly, Hill's N_2 allows a more intuitive comparison between communities (i.e., a community with $N_2 = 10$ species is twice as diverse as one with $N_2 = 5$) and is suitable for statistical analysis by basic linear regression methods[83]. Robust linear models of N_2 vs. time were regressed using the R package robustlmm.

Single-cycle viral growth kinetics

DF-1 or MDCK cells were seeded at $4x10^5$ cells per well in 6 well dishes 24 h prior to infection. GFHK99 WT / VAR_1 virus mixture was serially diluted using PBS. Synchronized infection conditions as described above were used with acid inactivation of inoculum virus and addition of ammonium chloride medium at 3 h post-infection. At each time point, 120 µl supernatant was collected. Viral titers for each sample were assessed by plaque assay in MDCK cells. Each MOI condition was used in 5-6 wells in parallel infections. Three wells served as technical replicates

for growth curve sampling while the remaining wells were harvested at 24 h post-infection to enumerate HA expressing cells via flow cytometry. In cases where acid inactivation was inefficient, the replicate was eliminated, and data are plotted in duplicate.

Effect of increasing multiple infection on viral RNA replication

For DF-1 and MDCK cell experiments, 12 well plates were seeded with $3x10^5$ cells per well 24 h prior to infection. For HTBE cells, cells were cultured at an air-liquid interface as previously described[69]. Cell surfaces were rinsed three times with PBS prior to inoculation. Triplicate wells were then mock infected with PBS or inoculated with 0.005 PFU per cell of GFHK99 WT and 0, 0.1, 0.5, 1, 3, and 5 PFU per cell of GFHK99 VAR_2 virus and placed at 37°C. After 55 minutes, inoculum was aspirated, and cells were rinsed three times with PBS and virus medium was added at 500 µl per well. Media was exchanged for ammonium chloride treated media 3 h later. At 12 h post infection, virus media was removed and cells were harvested using RNAprotect Cell Reagent (Qiagen). RNA was extracted using RNAeasy columns (Qiagen). RNA was diluted to 500 ng per µL for MDCK cells and 120 ng per µL for DF-1 cells. A 12 µl volume of this diluted RNA was used in reverse transcription with Maxima RT per protocol instructions. Digital droplet PCR was performed on the resultant cDNA using a combination of PB2, M, and NS primers specific for the GFHK99 WT virus (final primer concentration of 200 nM) with QX200™ ddPCR™ EvaGreen Supermix (Bio-Rad). WT copy number is determined as cDNA copies per ng of input RNA. WT fold-change was calculated by dividing the copies/ng result obtain in each VAR_2 positive condition by value of copies per ng from the average of the triplicate WT-only samples.

Strand-specific quantification of viral RNA species over time

MOIs used in this experiment were 0.5 RNA copies per cell for the low MOI and 3.0 HA expressing units/cell for the high MOI. Concentrations of virus mixtures in RNA copies per mL

were determined by quantifying at least four gene segments by ddPCR and taking the average. HA expressing units per mL was measured by counting HA positive cells via flow cytometry in the relevant cell type. Specifically, cells were infected with serial dilutions of virus under synchronized, single cycle conditions. At 24 h post infection, cells were harvested and flow cytometry was performed as described above, targeting His and HA epitope tags. HA expression units per mL for each virus and cell combination was calculated based on the linear range of $\%HA^+$ cells plotted as a function of volume of virus added to cells[84] (**Supplementary Figure 3**).

Viruses used for this experiment were the same GFHK99 WT/VAR_1 or MaMN99 WT/VAR virus mixtures used to measure reassortment, but in this case each mixture was considered as a single virus population (i.e. the RT ddPCR assay outlined below to quantify viral m/c/vRNA does not differentiate between WT and VAR genotypes).

Twelve well plates were seeded with $2x10^5$ cells per well of MDCK or DF-1 cells and incubated at 37°C for 24 h. Synchronized, single cycle infection conditions were used, as described above. Chilled virus was added at a volume of 125 µL per well. At 0, 1, 2, 4, 6, 8, and 10 h post infection, virus medium was aspirated and cells were harvested using 400 µL of CELLprotect solution (Qiagen). RNA was extracted from infected cells using the Qiagen RNAeasy Mini kit. Three reverse transcription reactions per sample were set up with three different primers, each containing different nucleotide barcode tags and targeting a distinct species (mRNA, vRNA, and cRNA) of segment 8 (**Supplementary Table 4**). Maxima RT was used according to the manufacturer's instructions and combined with 300 ng MDCK or 150 ng DF-1 RNA.

Absolute copy number of cDNA was determined by ddPCR. Forward and reverse primers for vRNA, mRNA, or cRNA of NS at a total concentration of 200 nM were combined with diluted cDNA and QX200™ ddPCR™ EvaGreen Supermix (Bio-Rad). Primer sequences are given in **Supplementary Table 4**. Thermocycler protocol was as follows: 95°C for 5 min, [95°C

for 30s, 57°C for 60s] repeat 40x, 4°C for 5 min, 90°C for 5 min, 4°C hold. Copy number was normalized to RNA input to give final results in units of copy number per ng RNA.

Single-cell mRNA sequencing

For this assay, viruses were titered in DF-1 cells using flow cytometry with anti-NP antibody (Abcam, clone 9G8). DF-1 cells were used because they give allow more sensitive detection of GFHK99 virus infection than MDCK cells. Cells were infected with serial dilutions of virus under synchronized, single-cycle conditions. At 24 h post-infection, cells were harvested and flow cytometry was performed as described above, targeting NP. HA expression units per mL for each virus and cell combination was calculated based on the linear range of % cells NP^+ plotted as a function of volume of virus added to cells[84] (**Supplementary Figure 3**).

To preform single-cell mRNA sequencing, MDCK and DF-1 cells were seeded into 6-well plates at $5x10^5$ cells per well. At 24 h post seeding, MDCK and DF-1 cells from an extra well were harvested and re-counted to ensure accuracy of cell number for infection. MDCK or DF-1 cells were then infected with a 1:1 ratio of GFHK99 WT virus and GFHK99 $mVAR_1$ virus that amounted to a MOI of 0.02, 0.06, 0.2 or 0.6 NP units per cell. GFHK99 $mVAR_2$ virus was added to MDCK and DF-1 cell infections at MOIs of 1 PFU per cell and 0.1 PFU per cell, respectively. Virus stocks were diluted serially in cold 1X PBS and incubated on ice until use. Before infection, cells were washed three times with cold 1X PBS and placed on ice. To infect, cells were inoculated with a 200 µL volume inoculum (at appropriate concentrations) and placed on ice for 45 minutes, with rocking every 10 minutes. The inoculum was then aspirated and 2 mL of pre-warmed (at 37ºC) virus medium was added. Plates were incubated at 37ºC for 3 h. Afterwards, the virus medium was replaced with 2 mL of pre-warmed virus medium supplemented with HEPES buffer and NH_4Cl at final concentrations of 50 mM and 20 mM, respectively. Plates were placed back into the incubator for an additional 5 h. Subsequently,

culture media was aspirated and cells washed once with 1X PBS. Cells were then trypsinized with 200 µL of 0.25% Trypsin EDTA until all cells came off the plate and were mono-dispersed. To each well, 0.5 mL of virus medium was added and replicates were pooled (2 wells per MOI). Cells for each sample were counted. Samples were spun at 150 rcf for 3 minutes and washed with 0.5 mL of 1X PBS/0.04% BSA. Washings were performed two more times. Finally, cells were resuspended with 1X PBS/0.04% BSA to get a final cell count of 7 x 10^5 cells per mL for each sample. Preparation for single-cell transcriptomic sequencing follows the protocol for 10x Genomics Chromium Single Cell platform.

Analysis of viral transcripts from single cells was performed with the sequencing data from all experiments in R using the CellRanger package (https://github.com/bpickett/Influenza-10X). Briefly, the CellRanger software assigns each read to individual cells and transcripts based on two sets of unique molecular identifiers that are ligated prior to amplification. This approach allows the quantification of amplification bias at both the cellular and transcript levels. The first step of the analytical workflow was to map the reads to concatenated transcriptomes of IAV with the transcriptomes of dog or chicken to analyze MDCK and DF-1 cell infections, respectively. Protein coding regions for the dog and chicken transcriptomes were identified in the GTF file associated with genome builds CanFam3.1.94 and Gallus gallus-5.0.94, respectively, while IAV coding regions were extracted from the reverse-complement sequences of the GFHK99 strains. For each experiment, all transcripts with non-zero numbers of mapped reads were then normalized to the median number of transcripts per cell to enable cross-experiment comparison. The read counts for all eight unspliced IAV transcripts for each MOI and cell type were subsequently extracted from the complete set and saved in separate files. A quantitative analysis was then performed to compare the number of IAV transcripts that were identified from each of the experimental variables. Unless otherwise stated, data was analyzed

using total viral transcripts, derived from all eight vRNA segments. The aligned sequencing data is available on the GEO database with the accession number GSE135553.

Single-cell sorting assay for measurement of P_P values

Segment specific P_P values were determined as previously described for influenza A/Panama/2007/99 (H3N2) virus[19], and as follows. $4x10^5$ MDCK cells were seeded into each well of a 6-well dish. 24 h later, cells were washed 3x with PBS and inoculated with 0.018 PFU per cell of GFHK99 WT virus and 1 PFU per cell of GFHK99 VAR_2 virus in a 250 µL volume of PBS. Virus was allowed to attach at 37°C for 1 h. Inoculum was then removed and cells were rinsed 3x with PBS and 2 mL of virus medium was added to the well. After 1 h at 37°C, medium was removed and cells were washed 3x with PBS and harvested by addition of Cell Dissociation Buffer (Corning). Cells were resuspended in complete medium and washed 3x with 2 mL FACS buffer (2% FBS in PBS). A final resuspension step was performed in PBS containing 1% FBS, 10 mM HEPES, and 0.1% EDTA. Cells were strained through a cell strainer cap (Falcon) and sorted on a BD Aria II cell sorter. Gating was performed to remove debris and multiplets and one event per well was sorted into each well of a 96 well plate containing MDCK monolayers at 30% confluency in 50 µl virus medium supplemented with 1 ug per mL TPCK-treated trypsin. Following the sort, an additional 50 µl of virus medium plus trypsin was added to each well and plates were centrifuged at 1,800 rpm for 2 minutes to promote cell attachment. Plates were incubated at 37°C for 48 h to allow propagation of virus from the sorted cell.

RNA was extracted from infected cells in the 96 well plate using a ZR-96 Viral RNA Kit (Zymo Research) per manufacturer instructions. Extracted RNA was converted to cDNA using universal influenza primers[85] and Maxima RT according to manufacturer instructions. After conversion, cDNA was diluted 1:4 with nuclease-free water and used as template (4 µL per reaction) for segment-specific qPCR using SsoFast EvaGreen Supermix (Bio-Rad) in 10 µl

reactions, with 200 nM final primer concentration. Primers employed targeted each segment of GFHK99 WT virus, as well as the PB2 and PB1 segments of GFHK99 VAR_2 virus. Primer sequences are listed in **Supplementary Table 3**.

Given the MOI of GFHK99 WT virus used in the experiments, an appreciable number of wells are expected to receive two or more viral genomes, and so a mathematical adjustment is needed to estimate the probability of each genome segment being delivered by a single virion. Using the relationship between MOI and the fraction of cells infected from Poisson statistics, i.e., $f = 1 - e^{-\mathrm{MOI}}$, the probability of the *i*th segment being present in a singly infected cell, or $P_{P,i}$ can be calculated from the 96-well plate using the following equation:

$$P_{P,i} = \frac{\mathrm{MOI}_i}{\mathrm{MOI}_{\mathrm{wt}}} = \frac{-\ln\,(1 - f_i)}{-\ln\,(1 - f_{\mathrm{wt}})} = \frac{\ln\,(1 - \frac{C_i}{A})}{\ln\,(1 - \frac{B}{A})}$$

where A is the number of VAR_2^+ wells, B is the number of WT^+ wells (containing any WT segment), and C_i is the number of wells positive for the WT segment in question. Wells that were negative for VAR_2 virus segments were excluded from analysis.

794 **Figure Legends**

795 **Figure 1. Coinfection and reassortment frequencies indicate that IAV multiplicity**
796 **dependence varies with virus strain and host species.** A-D) MDCK or DF-1 cells were
797 coinfected with homologous WT and VAR viruses of either GFHK99 or MaMN99 strain
798 backgrounds at a range of MOIs. Following a single cycle of infection, cells were analyzed for
799 HA expression by flow cytometry and plaque clones derived from cell supernatants were
800 genotyped. The relationship between % cells HA positive and % cells dually HA positive (A)
801 varies with strain and cell type, resulting in curves of differing % linearity (B). GFHK99 and
802 MaMN99 viruses exhibit different reassortment levels in MDCK cells, but both show high
803 reassortment relative to a theoretical prediction in which singly infected and multiply infected
804 cells have equivalent burst sizes (C). GFHK99 virus reassortment levels differ in MDCK and DF-
805 1 cells, but reassortment under both conditions remains high relative to the theoretical prediction
806 in which multiple infection confers no advantage (D). In guinea pigs (n=6), GFHK99 WT and
807 VAR_1 viruses exhibit higher reassortment than MaMN99 WT and VAR viruses, as indicated by
808 increased genotypic diversity (E). The GFHK99 WT and VAR_1 viruses exhibit higher
809 reassortment in guinea pigs than in quail (n=5) (F). Guinea pig data shown in panels E and F
810 are the same. Shading represents 95% CI.

811

812 **Figure 2. Increasing MOI increases viral productivity at sub-saturating, but not saturating**
813 **MOIs.** MDCK and DF-1 cells were infected under single cycle conditions at a range of MOIs in
814 triplicate wells for each MOI. A-E) Viral titers observed at the indicated MOIs are plotted against
815 time post-infection. F) Fold change in amplification (viral input / maximum output) relative to the
816 MOI=0.01 PFU per cell condition is plotted for each virus-cell pairing. G) Burst size, calculated
817 as maximum PFU output / number of HA^+ cells detected by flow cytometry, is plotted for each
818 virus-cell pairing tested in the higher MOI range.

819

820 **Figure 3. Coinfection and reassortment of chimeric viruses reveals a major role for the**
821 **viral polymerase.** Reverse genetics was used to place one or more genes from GFHK99 virus
822 into a MaMN99 background. Coinfections with homologous WT and VAR strains were
823 performed in MDCK cells as in Figure 1. The relationship between % cells HA positive and %
824 cells dually HA positive (A) varies with genotype, resulting in curves of differing % linearity (B).
825 Reassortment levels vary with genotype (C), with the chimeric strain carrying GFHK99 PB2,
826 PB1, PA and NP segments exhibiting comparable levels to GFHK99. Experimental results are
827 compared to a theoretical prediction in which singly infected and multiply infected cells have
828 equivalent burst sizes (Prediction). Differences in reassortment levels among the viruses tested
829 are highlighted by plotting the % reassortment at 10% HA^+ cells, as interpolated from each
830 regression curve (D). Data shown for GFHK99 and MaMN99 viruses are the same as those
831 displayed in Figure 1.
832
833 **Figure 4. Coinfection enhances GFHK99 vRNA synthesis in a dose and host dependent**
834 **manner.** Cells were coinfected with 0.005 PFU per cell of GFHK99 WT virus and increasing
835 doses of GFHK99 VAR_2 virus. A) In MDCK and DF-1 cells, the fold change in WT vRNA copy
836 number, relative to that detected in the absence of GFHK99 VAR_2 virus, is plotted for various
837 doses of GFHK99 VAR_2 virus. B) In HTBE cells, the fold change in WT vRNA copy number,
838 relative to that detected in the absence of GFHK99 VAR_2 virus, is plotted for various doses of
839 GFHK99 VAR_2 virus. n=3 cell culture dishes per condition. Error bars represent standard error.
840
841 **Figure 5. High multiplicity of infection is needed for robust GFHK99 polymerase activity**
842 **in MDCK cells.** Dishes of MDCK or DF-1 cells (n=3) were infected with GFHK99 or MaMN99
843 virus at low (0.5 RNA copies per cell) or high (3 HA expressing units per cell) MOI. NS segment
844 vRNA, mRNA, and cRNA were quantified at the indicate time points (A-F). The average fold
845 change from initial (t=0) to peak RNA copy number is plotted for low MOI infections (G) and high

MOI infections (H). Error bars represent standard error. Significance was assessed by two-way ANOVA with Dunnett's test for multiple comparisons: *p < 0.05, **<0.01, ***<0.001. ns = not significant.

Figure 6. GFHK99 viral transcription is uniformly low in MDCK cells in the absence of coinfecting virus. (A) DF-1 or MDCK cells were infected with GFHK99 WT virus at three different MOIs (0.67, 0.2, 0.6 NP units per cell), and the transcriptomes of 1,816 individual infected cells were elucidated using the 10X Genomics Chromium platform. Ridge plots show distributions of $\log_{10}$-transformed viral mRNA abundance, for all eight viral transcripts combined, in individual infected cells. The data are stratified by cell type (MDCK cells in blue, DF-1 cells in pink), MOI, and the presence of polymerase complex (light shading = cells missing PB2, PB1, PA, or NP; dark shading = cells in which PB2, PB1, and PA are all detected). The absence of a dark shaded distribution for MDCK cells at the lowest MOI is due to the absence of any cells in which all four of these segments were detected. (B) DF-1 or MDCK cells were infected and sequenced as in (A), but the inocula contained a 1:1 mixture of GFHK99 WT and GFHK99 $mVAR_1$ viruses at three different total MOIs (0.67, 0.2, 0.6 NP units per cell) and a constant amount of GFHK99 $mVAR_2$ virus (0.1 PFU per cell in DF-1 cells, 1.0 PFU per cell in MDCK cells). Left facet shows data from (A), and right facet shows data from WT/$mVAR_1$ coinoculations with $mVAR_2$ virus, with WT and $mVAR_1$ transcript abundances partitioned into separate distributions in each infection. Vertical lines denote the median of each distribution. UMI = unique molecular identifier.

Figure 7. Incomplete GFHK99 virus genomes are present in MDCK cells but not sufficiently abundant to account for observed reassortment. Incomplete viral genomes were quantified experimentally by a single-cell based assay which relies on the amplification of incomplete viral genomes of GFHK99 WT virus (0.018 PFU per cell) by a genetically similar

872 coinfecing virus, GFHK99 VAR_2. Based on the rate of detection of GFHK99 WT virus segments
873 in this assay, the probability that a given segment would be present and replicated in a singly
874 infected MDCK cell is reported as P_P. A) Summary of experimental P_P data. n = 2 biological
875 replicates, shown in blue and red. Shading represents 95% CI. B) Experimentally obtained P_P
876 values in A were used to parameterize a computational model[18]. Levels of reassortment
877 predicted using the experimentally determined parameters are shown in red and blue. Levels of
878 reassortment predicted if P_P=1.0 are shown with the dashed line. Observed reassortment of
879 GFHK99 WT and VAR viruses in MDCK cells are shown with black circles. Observed data are
880 the same as those plotted in Figure 1.

881

882 **References**

883 1 Leeks, A., Sanjuan, R. & West, S. A. The evolution of collective infectious units in viruses. *Virus*
884 *Res* **265**, 94-101, doi:10.1016/j.virusres.2019.03.013 (2019).
885 2 Brooke, C. B. Population Diversity and Collective Interactions during Influenza Virus Infection. *J*
886 *Virol* **91**, doi:10.1128/JVI.01164-17 (2017).
887 3 Sanjuan, R. Collective Infectious Units in Viruses. *Trends Microbiol* **25**, 402-412,
888 doi:10.1016/j.tim.2017.02.003 (2017).
889 4 Santiana, M. *et al.* Vesicle-Cloaked Virus Clusters Are Optimal Units for Inter-organismal Viral
890 Transmission. *Cell Host Microbe* **24**, 208-220 e208, doi:10.1016/j.chom.2018.07.006 (2018).
891 5 Chen, Y. H. *et al.* Phosphatidylserine vesicles enable efficient en bloc transmission of
892 enteroviruses. *Cell* **160**, 619-630, doi:10.1016/j.cell.2015.01.032 (2015).
893 6 Erickson, A. K. *et al.* Bacteria Facilitate Enteric Virus Co-infection of Mammalian Cells and
894 Promote Genetic Recombination. *Cell Host Microbe* **23**, 77-88 e75,
895 doi:10.1016/j.chom.2017.11.007 (2018).
896 7 Andreu-Moreno, I. & Sanjuan, R. Collective Infection of Cells by Viral Aggregates Promotes Early
897 Viral Proliferation and Reveals a Cellular-Level Allee Effect. *Current biology : CB* **28**, 3212-
898 3219.e3214, doi:10.1016/j.cub.2018.08.028 (2018).
899 8 Cuevas, J. M., Duran-Moreno, M. & Sanjuan, R. Multi-virion infectious units arise from free viral
900 particles in an enveloped virus. *Nat Microbiol* **2**, 17078, doi:10.1038/nmicrobiol.2017.78 (2017).
901 9 Cifuentes-Munoz, N., Dutch, R. E. & Cattaneo, R. Direct cell-to-cell transmission of respiratory
902 viruses: The fast lanes. *PLoS Pathog* **14**, e1007015, doi:10.1371/journal.ppat.1007015 (2018).
903 10 Lawrence, D. M. *et al.* Measles virus spread between neurons requires cell contact but not CD46
904 expression, syncytium formation, or extracellular virus production. *J Virol* **74**, 1908-1918,
905 doi:10.1128/jvi.74.4.1908-1918.2000 (2000).
906 11 McDonald, D. *et al.* Recruitment of HIV and its receptors to dendritic cell-T cell junctions. *Science*
907 **300**, 1295-1297, doi:10.1126/science.1084238 (2003).
908 12 Shirogane, Y., Watanabe, S. & Yanagi, Y. Cooperation between different RNA virus genomes
909 produces a new phenotype. *Nat Commun* **3**, 1235, doi:10.1038/ncomms2252 (2012).

13 Ciota, A. T., Ehrbar, D. J., Van Slyke, G. A., Willsey, G. G. & Kramer, L. D. Cooperative interactions in the West Nile virus mutant swarm. *BMC Evol Biol* **12**, 58, doi:10.1186/1471-2148-12-58 (2012).

14 Xue, K. S., Hooper, K. A., Ollodart, A. R., Dingens, A. S. & Bloom, J. D. Cooperation between distinct viral variants promotes growth of H3N2 influenza in cell culture. *Elife* **5**, e13974, doi:10.7554/eLife.13974 (2016).

15 Brooke, C. B., Ince, W. L., Wei, J., Bennink, J. R. & Yewdell, J. W. Influenza A virus nucleoprotein selectively decreases neuraminidase gene-segment packaging while enhancing viral fitness and transmissibility. *Proc Natl Acad Sci U S A* **111**, 16854-16859, doi:10.1073/pnas.1415396111 (2014).

16 Brooke, C. B. *et al.* Most influenza a virions fail to express at least one essential viral protein. *J Virol* **87**, 3155-3162, doi:10.1128/jvi.02284-12 (2013).

17 Sun, J. & Brooke, C. B. Influenza A Virus Superinfection Potential Is Regulated by Viral Genomic Heterogeneity. *MBio* **9**, doi:10.1128/mBio.01761-18 (2018).

18 Fonville, J. M., Marshall, N., Tao, H., Steel, J. & Lowen, A. C. Influenza Virus Reassortment Is Enhanced by Semi-infectious Particles but Can Be Suppressed by Defective Interfering Particles. *PLoS Pathog* **11**, e1005204, doi:10.1371/journal.ppat.1005204 (2015).

19 Jacobs, N. T. *et al.* Incomplete influenza A virus genomes occur frequently but are readiliy complemented during localized viral spread. *Nature Communications* (2019).

20 Nayak, D. P. Defective interfering influenza viruses. *Annu Rev Microbiol* **34**, 619-644, doi:10.1146/annurev.mi.34.100180.003155 (1980).

21 Von Magnus, P. Incomplete forms of influenza virus. *Adv Virus Res* **2**, 59-79 (1954).

22 Brooke, C. B. Biological activities of 'noninfectious' influenza A virus particles. *Future Virol* **9**, 41-51, doi:10.2217/fvl.13.118 (2014).

23 Timm, C., Gupta, A. & Yin, J. Robust kinetics of an RNA virus: Transcription rates are set by genome levels. *Biotechnology and bioengineering* **112**, 1655-1662, doi:10.1002/bit.25578 (2015).

24 Boulle, M. *et al.* HIV Cell-to-Cell Spread Results in Earlier Onset of Viral Gene Expression by Multiple Infections per Cell. *PLoS Pathog* **12**, e1005964, doi:10.1371/journal.ppat.1005964 (2016).

25 Sanjuán, R. & Thoulouze, M.-I. Why viruses sometimes disperse in groups?(†). *Virus Evol* **5**, vez014-vez014, doi:10.1093/ve/vez014 (2019).

26 Sigal, A. *et al.* Cell-to-cell spread of HIV permits ongoing replication despite antiretroviral therapy. *Nature* **477**, 95-98, doi:10.1038/nature10347 (2011).

27 Webster, R. G., Hinshaw, V. S., Bean, W. J., Jr., Turner, B. & Shortridge, K. F. Influenza viruses from avian and porcine sources and their possible role in the origin of human pandemic strains. *Dev Biol Stand* **39**, 461-468 (1977).

28 Wright, P. F., Neumann, G. & Kawaoka, Y. in *Fields Virology* Vol. 1 (ed D. M. & Howley Knipe, P. M.) 1691-1740 (Lippincott-Raven, 2006).

29 Webster, R. G., Bean, W. J., Gorman, O. T., Chambers, T. M. & Kawaoka, Y. Evolution and ecology of influenza A viruses. *Microbiol Rev* **56**, 152-179 (1992).

30 Webster, R. G., Shortridge, K. F. & Kawaoka, Y. Influenza: interspecies transmission and emergence of new pandemics. *FEMS Immunol Med Microbiol* **18**, 275-279, doi:S0928-8244(97)00058-8 [pii] (1997).

31 Taubenberger, J. K. & Morens, D. M. 1918 Influenza: the mother of all pandemics. *Emerg Infect Dis* **12**, 15-22 (2006).

32 Viboud, C., Miller, M., Olson, D., Osterholm, M. & Simonsen, L. Preliminary Estimates of Mortality and Years of Life Lost Associated with the 2009 A/H1N1 Pandemic in the US and

Comparison with Past Influenza Seasons. *PLoS Curr*, RRN1153, doi:k/-/-/35hpbywfdwl4n/8 [pii] (2010).

33 Long, J. S., Mistry, B., Haslam, S. M. & Barclay, W. S. Host and viral determinants of influenza A virus species specificity. *Nat Rev Microbiol* **17**, 67-81, doi:10.1038/s41579-018-0115-z (2019).

34 Rajsbaum, R. *et al.* Species-specific inhibition of RIG-I ubiquitination and IFN induction by the influenza A virus NS1 protein. *PLoS Pathog* **8**, e1003059, doi:10.1371/journal.ppat.1003059 (2012).

35 Manz, B. *et al.* Pandemic influenza A viruses escape from restriction by human MxA through adaptive mutations in the nucleoprotein. *PLoS Pathog* **9**, e1003279, doi:10.1371/journal.ppat.1003279

PPATHOGENS-D-12-02051 [pii] (2013).

36 Lowen, A. C. Constraints, Drivers, and Implications of Influenza A Virus Reassortment. *Annual review of virology* **4**, 105-121, doi:10.1146/annurev-virology-101416-041726 (2017).

37 Scholtissek, C. Molecular evolution of influenza viruses. *Virus Genes* **11**, 209-215 (1995).

38 Ma, E. J., Hill, N. J., Zabilansky, J., Yuan, K. & Runstadler, J. A. Reticulate evolution is favored in influenza niche switching. *Proc Natl Acad Sci U S A* **113**, 5335-5339, doi:10.1073/pnas.1522921113 (2016).

39 Kilbourne, E. D. Influenza pandemics of the 20th century. *Emerg Infect Dis* **12**, 9-14 (2006).

40 Smith, G. J. *et al.* Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* **459**, 1122-1125, doi:nature08182 [pii]

10.1038/nature08182 (2009).

41 Kayed, A. S. *et al.* Surveillance for avian influenza viruses in wild birds at live bird markets, Egypt, 2014-2016. *Influenza Other Respir Viruses* **13**, 407-414, doi:10.1111/irv.12634 (2019).

42 Xu, K. M. *et al.* The genesis and evolution of H9N2 influenza viruses in poultry from southern China, 2000 to 2005. *J Virol* **81**, 10389-10401, doi:10.1128/JVI.00979-07 (2007).

43 Russell, A. B., Trapnell, C. & Bloom, J. D. Extreme heterogeneity of influenza virus infection in single cells. *Elife* **7**, doi:10.7554/eLife.32303 (2018).

44 Zheng, G. X. *et al.* Massively parallel digital transcriptional profiling of single cells. *Nat Commun* **8**, 14049, doi:10.1038/ncomms14049 (2017).

45 Palese, P. & Shaw, M. L. in *Fields Virology* Vol. 1 (ed D. M. & Howley Knipe, P. M.) Ch. 41, 1151-1185 (Lippincott-Raven, 2013).

46 Lowen, A. C. It's in the mix: Reassortment of segmented viral genomes. *PLoS Pathog* **14**, e1007200, doi:10.1371/journal.ppat.1007200 (2018).

47 Krystal, M., Li, R., Lyles, D., Pavlakis, G. & Palese, P. Expression of the three influenza virus polymerase proteins in a single cell allows growth complementation of viral mutants. *Proc Natl Acad Sci U S A* **83**, 2709-2713, doi:10.1073/pnas.83.8.2709 (1986).

48 Naffakh, N., Tomoiu, A., Rameix-Welti, M. A. & van der Werf, S. Host restriction of avian influenza viruses at the level of the ribonucleoproteins. *Annu Rev Microbiol* **62**, 403-424, doi:10.1146/annurev.micro.62.081307.162746 (2008).

49 Mehle, A. & Doudna, J. A. Adaptive strategies of the influenza virus polymerase for replication in humans. *Proc Natl Acad Sci U S A* **106**, 21312-21316, doi:0911915106 [pii]

10.1073/pnas.0911915106 (2009).

50 Long, J. S. *et al.* Species difference in ANP32A underlies influenza A virus polymerase host restriction. *Nature* **529**, 101-104, doi:10.1038/nature16474 (2016).

51 Te Velthuis, A. J. W. *et al.* Mini viral RNAs act as innate immune agonists during influenza virus infection. *Nature microbiology* **3**, 1234-1242, doi:10.1038/s41564-018-0240-5 (2018).

52 Schelker, M. *et al.* Viral RNA Degradation and Diffusion Act as a Bottleneck for the Influenza A Virus Infection Efficiency. *PLOS Computational Biology* **12**, e1005075, doi:10.1371/journal.pcbi.1005075 (2016).

53 Butt, K. M. *et al.* Human infection with an avian H9N2 influenza A virus in Hong Kong in 2003. *J Clin Microbiol* **43**, 5760-5767, doi:43/11/5760 [pii]

10.1128/JCM.43.11.5760-5767.2005 (2005).

54 Peiris, M. *et al.* Human infection with influenza H9N2. *Lancet* **354**, 916-917, doi:S0140673699033115 [pii] (1999).

55 Guan, Y. *et al.* H9N2 influenza viruses possessing H5N1-like internal genomes continue to circulate in poultry in southeastern China. *J Virol* **74**, 9372-9380 (2000).

56 Guan, Y., Shortridge, K. F., Krauss, S. & Webster, R. G. Molecular characterization of H9N2 influenza viruses: were they the donors of the "internal" genes of H5N1 viruses in Hong Kong? *Proc Natl Acad Sci U S A* **96**, 9363-9367 (1999).

57 Lam, T. T. *et al.* The genesis and source of the H7N9 influenza viruses causing human infections in China. *Nature* **502**, 241-244, doi:10.1038/nature12515

nature12515 [pii] (2013).

58 Wu, A. *et al.* Sequential Reassortments Underlie Diverse Influenza H7N9 Genotypes in China. *Cell Host Microbe* **4**, 446-452, doi:S1931-3128(13)00298-9 [pii]

10.1016/j.chom.2013.09.001 (2013).

59 Gu, M., Xu, L., Wang, X. & Liu, X. Current situation of H9N2 subtype avian influenza in China. *Vet Res* **48**, 49-49, doi:10.1186/s13567-017-0453-2 (2017).

60 Pusch, E. A. & Suarez, D. L. The Multifaceted Zoonotic Risk of H9N2 Avian Influenza. *Vet Sci* **5**, 82, doi:10.3390/vetsci5040082 (2018).

61 White, M. C. & Lowen, A. C. Implications of segment mismatch for influenza A virus evolution. *J Gen Virol* **99**, 3-16, doi:10.1099/jgv.0.000989 (2018).

62 Villa, M. & Lassig, M. Fitness cost of reassortment in human influenza. *PLoS Pathog* **13**, e1006685, doi:10.1371/journal.ppat.1006685 (2017).

63 Varble, A. *et al.* Influenza A virus transmission bottlenecks are defined by infection route and recipient host. *Cell Host Microbe* **16**, 691-700, doi:10.1016/j.chom.2014.09.020 (2014).

64 Gallagher, M. E., Brooke, C. B., Ke, R. & Koelle, K. Causes and Consequences of Spatial Within-Host Viral Spread. *Viruses* **10**, doi:10.3390/v10110627 (2018).

65 Chao, L., Tran, T. & Matthews, C. Muller's Ratchet and the Advantage of Sex in the Rna Virus 6. *Evolution; international journal of organic evolution* **46**, 289-299, doi:10.1111/j.1558-5646.1992.tb02038.x (1992).

66 Froissart, R. *et al.* Co-infection weakens selection against epistatic mutations in RNA viruses. *Genetics* **168**, 9-19, doi:10.1534/genetics.104.030205 (2004).

67 Novella, I. S., Reissig, D. D. & Wilke, C. O. Density-dependent selection in vesicular stomatitis virus. *J Virol* **78**, 5799-5804, doi:10.1128/JVI.78.11.5799-5804.2004 (2004).

68 Wilke, C. O. & Novella, I. S. Phenotypic mixing and hiding may contribute to memory in viral quasispecies. *BMC Microbiol* **3**, 11, doi:10.1186/1471-2180-3-11 (2003).

69 Danzy, S. *et al.* Mutations to PB2 and NP proteins of an avian influenza virus combine to confer efficient growth in primary human respiratory cells. *J Virol* **88**, 13436-13446, doi:10.1128/JVI.01093-14 (2014).

70 Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. & Webster, R. G. A DNA transfection system for generation of influenza A virus from eight plasmids. **97**, 6108-6113, doi:10.1073/pnas.100133697 %J Proceedings of the National Academy of Sciences (2000).

1050 71 Schwartz, S. L. & Lowen, A. C. Droplet digital PCR: A novel method for detection of influenza
1051 virus defective interfering particles. *J Virol Methods* **237**, 159-165,
1052 doi:10.1016/j.jviromet.2016.08.023 (2016).
1053 72 Perez, D. R., Webby, R. J., Hoffmann, E. & Webster, R. G. Land-based birds as potential
1054 disseminators of avian mammalian reassortant influenza A viruses. *Avian Dis* **47**, 1114-1117
1055 (2003).
1056 73 Song, H., Nieto, G. R. & Perez, D. R. A new generation of modified live-attenuated avian
1057 influenza viruses using a two-strategy combination as potential vaccine candidates. *J Virol* **81**,
1058 9238-9248, doi:JVI.00893-07 [pii]

1059 10.1128/JVI.00893-07 (2007).
1060 74 Sorrell, E. M., Wan, H., Araya, Y., Song, H. & Perez, D. R. Minimal molecular constraints for
1061 respiratory droplet transmission of an avian-human H9N2 influenza A virus. *Proc Natl Acad Sci U*
1062 *S A* **106**, 7565-7570, doi:0900877106 [pii]

1063 10.1073/pnas.0900877106 (2009).
1064 75 Brown, J. D. *et al.* Intestinal excretion of a wild bird-origin H3N8 low pathogenic avian influenza
1065 virus in mallards (Anas Platyrhynchos). *Journal of wildlife diseases* **48**, 991-998,
1066 doi:10.7589/2011-09-280 (2012).
1067 76 Chen, H. *et al.* Partial and full PCR-based reverse genetics strategy for influenza viruses. *PLoS*
1068 *One* **7**, e46378, doi:10.1371/journal.pone.0046378

1069 PONE-D-12-21490 [pii] (2012).
1070 77 Marshall, N., Priyamvada, L., Ende, Z., Steel, J. & Lowen, A. C. Influenza virus reassortment
1071 occurs with high frequency in the absence of segment mismatch. *PLoS Pathog* **9**, e1003421,
1072 doi:10.1371/journal.ppat.1003421 (2013).
1073 78 Reed, L. J. & Muench, H. A SIMPLE METHOD OF ESTIMATING FIFTY PER CENT ENDPOINTS12.
1074 *American Journal of Epidemiology* **27**, 493-497, doi:10.1093/oxfordjournals.aje.a118408 (1938).
1075 79 Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G. & Pryor, R. J. High-resolution
1076 genotyping by amplicon melting analysis using LCGreen. *Clin Chem* **49**, 853-860 (2003).
1077 80 Richard, M., Herfst, S., Tao, H., Jacobs, N. T. & Lowen, A. C. Influenza A virus reassortment is
1078 limited by anatomical compartmentalization following co-infection via distinct routes. *J Virol*,
1079 doi:10.1128/JVI.02063-17 (2017).
1080 81 Simpson, E. H. Measurement of diversity. *Nature* **163**, 688-688 (1949).
1081 82 Hill, M. O. Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**, 427-
1082 432 (1973).
1083 83 Jost, L. Entropy and diversity. *Oikos* **113**, 363-375 (2006).
1084 84 Drayman, N., Kler, S., Ben-nun-Shaul, O. & Oppenheim, A. Rapid method for SV40 titration. *J*
1085 *Virol Methods* **164**, 145-147, doi:10.1016/j.jviromet.2009.12.003 (2010).
1086 85 Zhou, B. *et al.* Single-reaction genomic amplification accelerates sequencing and vaccine
1087 production for classical and Swine origin human influenza a viruses. *J Virol* **83**, 10309-10313,
1088 doi:JVI.01109-09 [pii]

1089 10.1128/JVI.01109-09 (2009).

1090

Figure 1

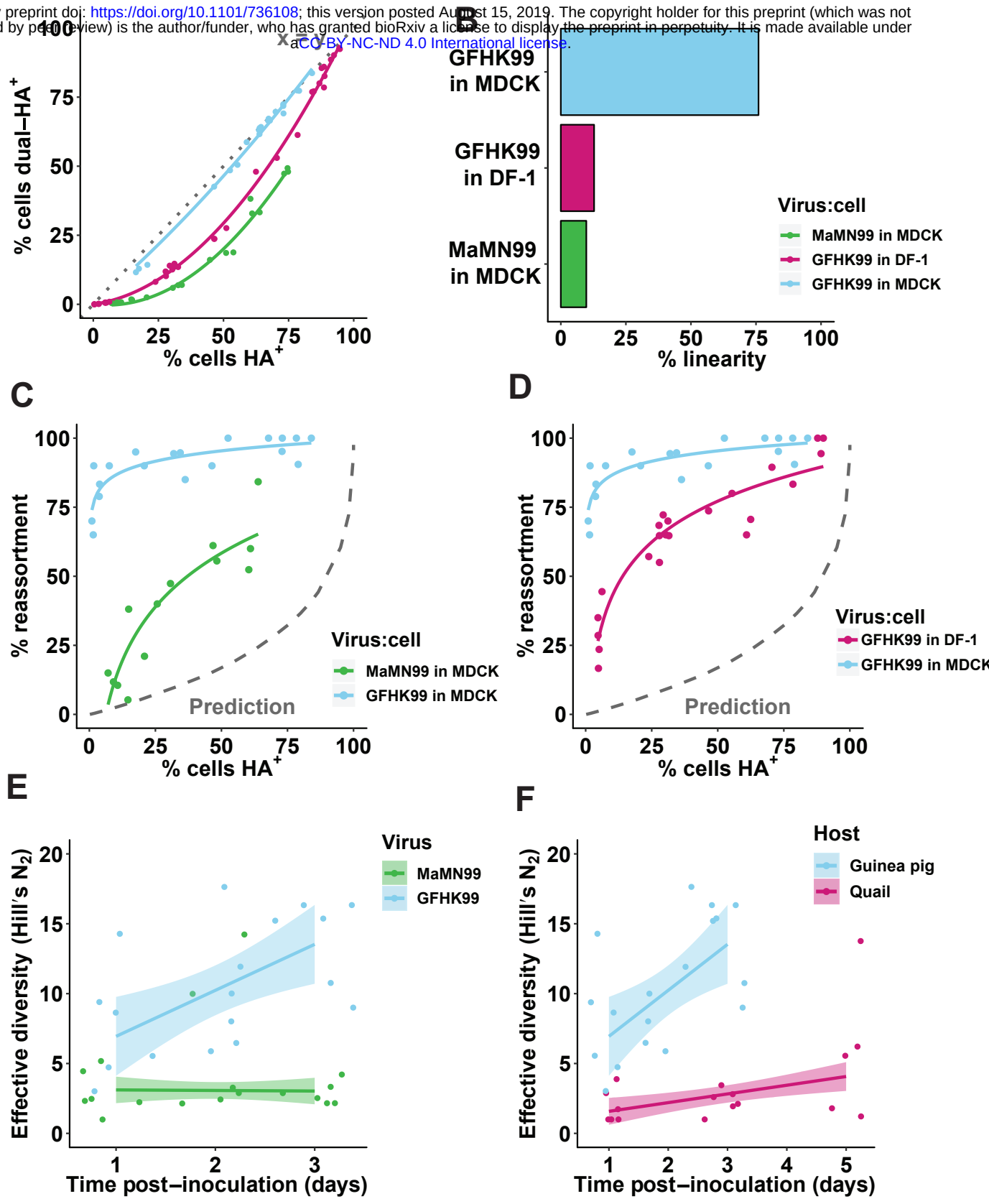

Figure 2

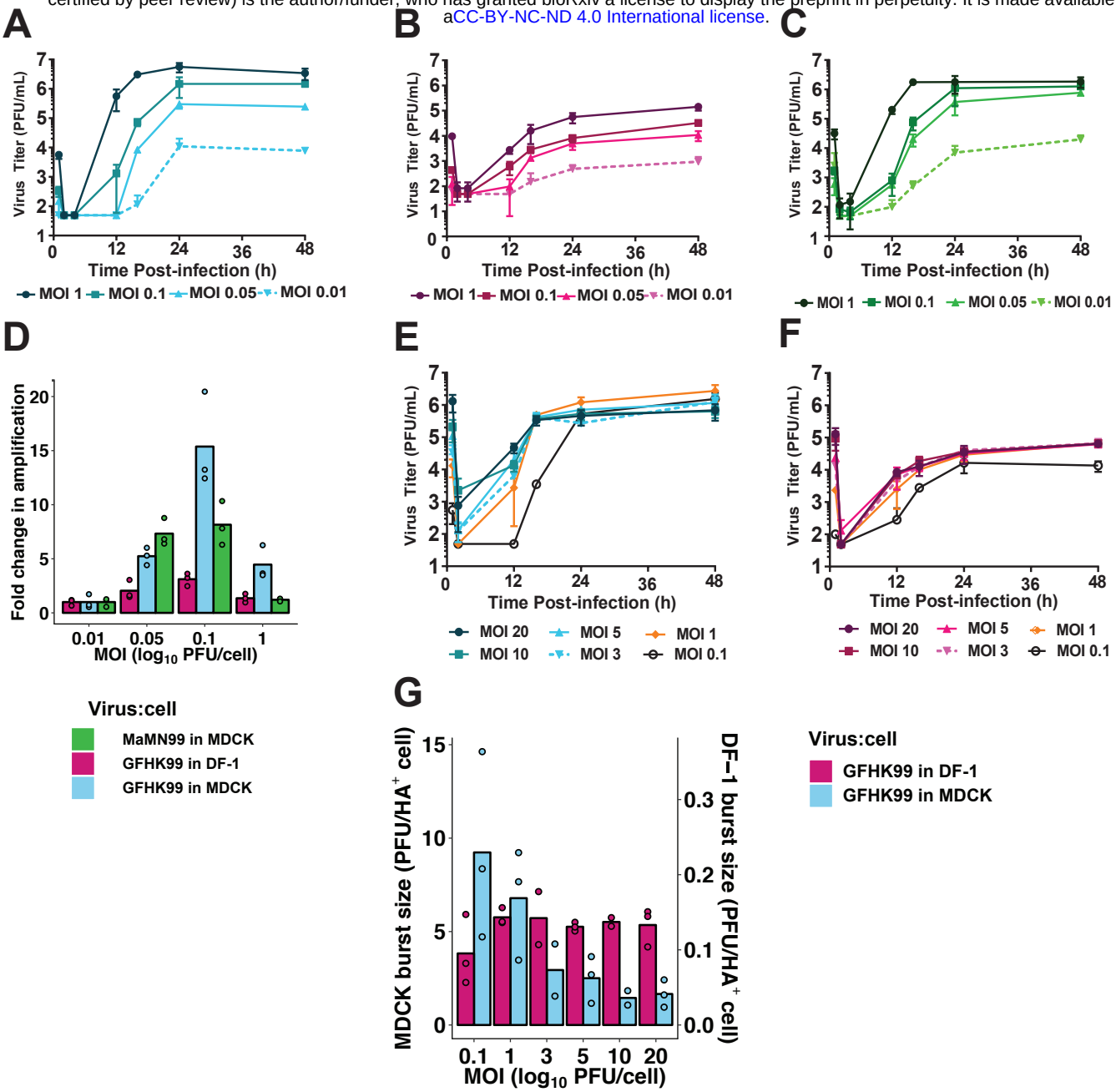

Figure 3

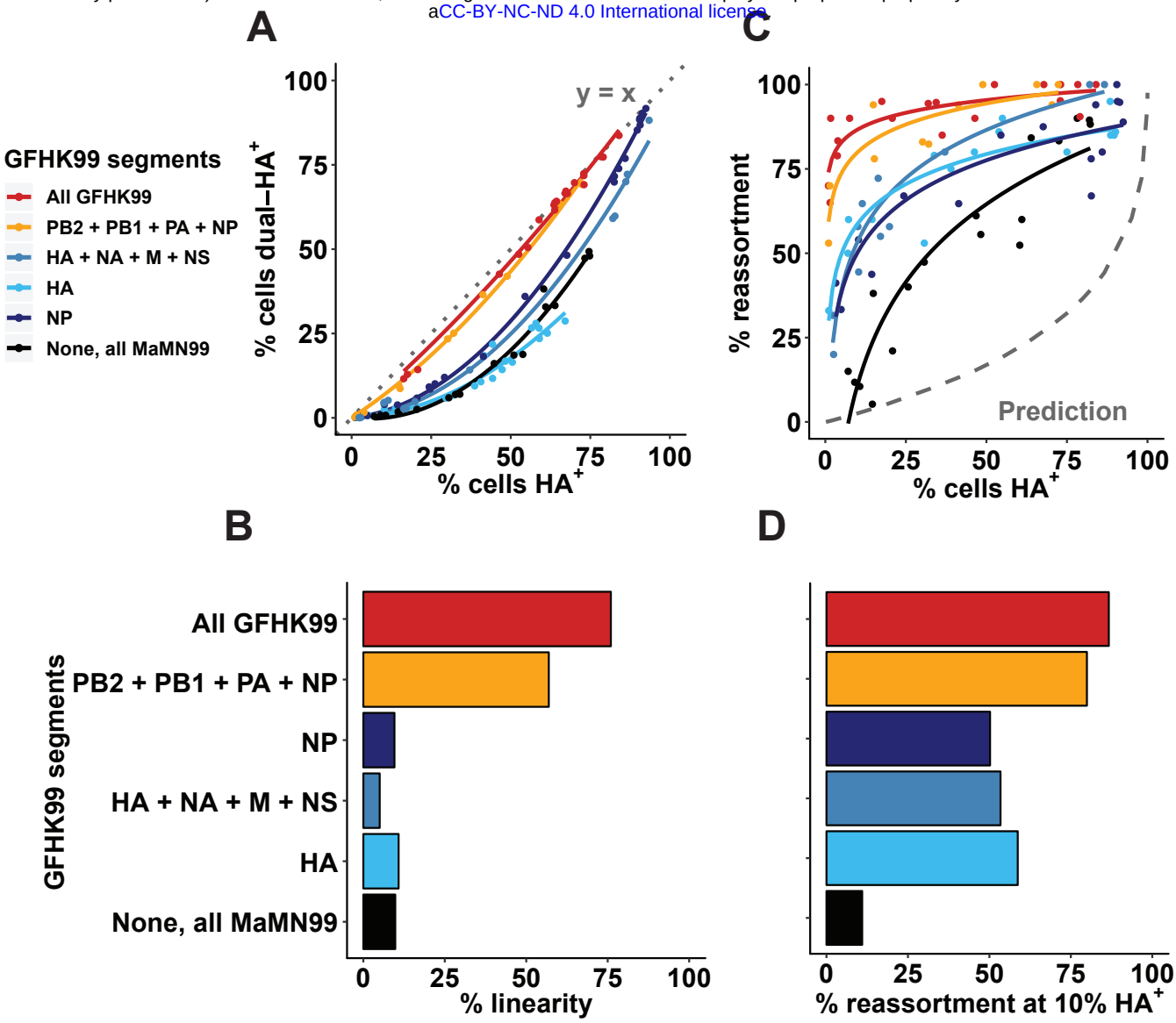

Figure 4

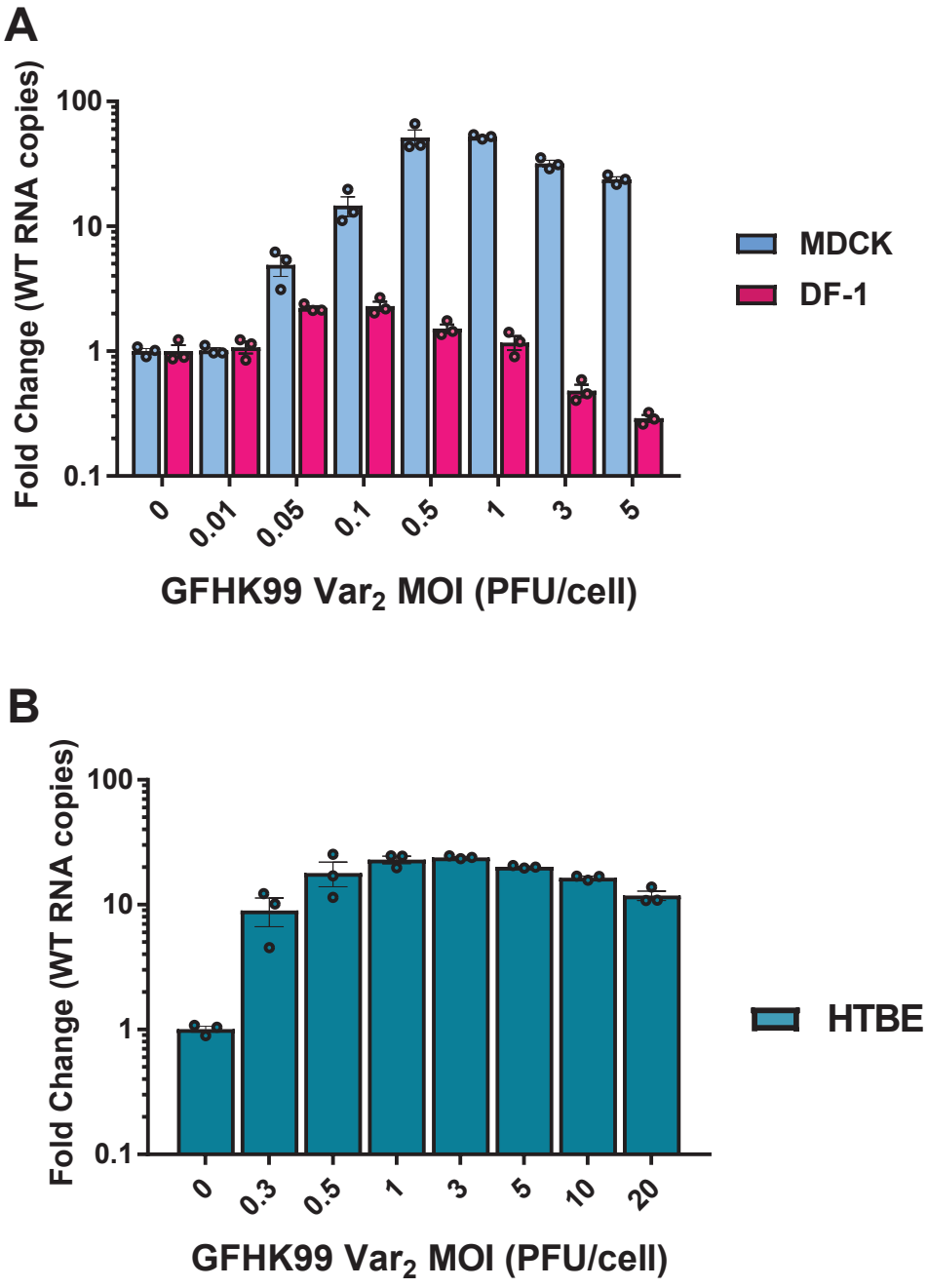

Figure 5

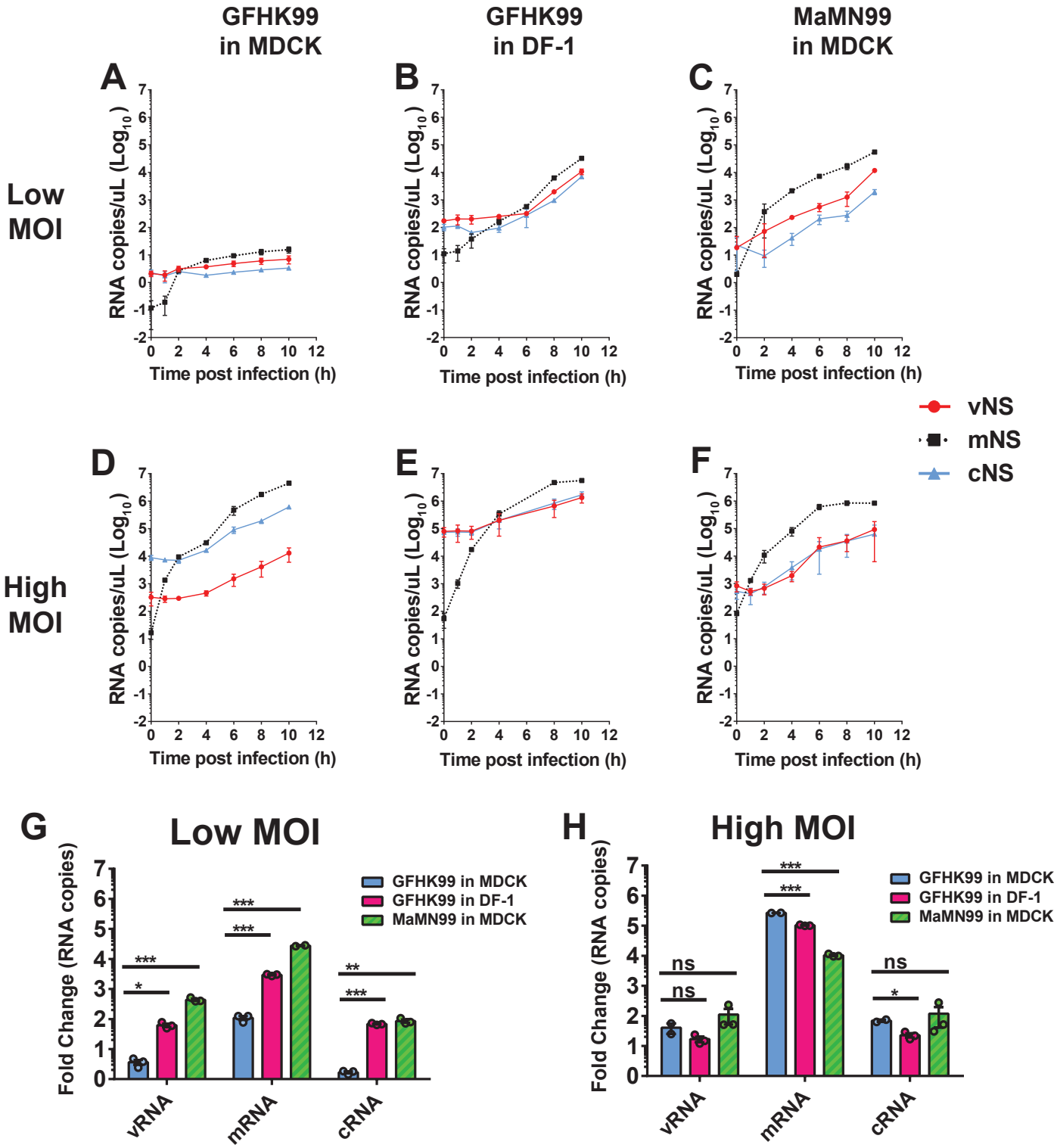

Figure 6

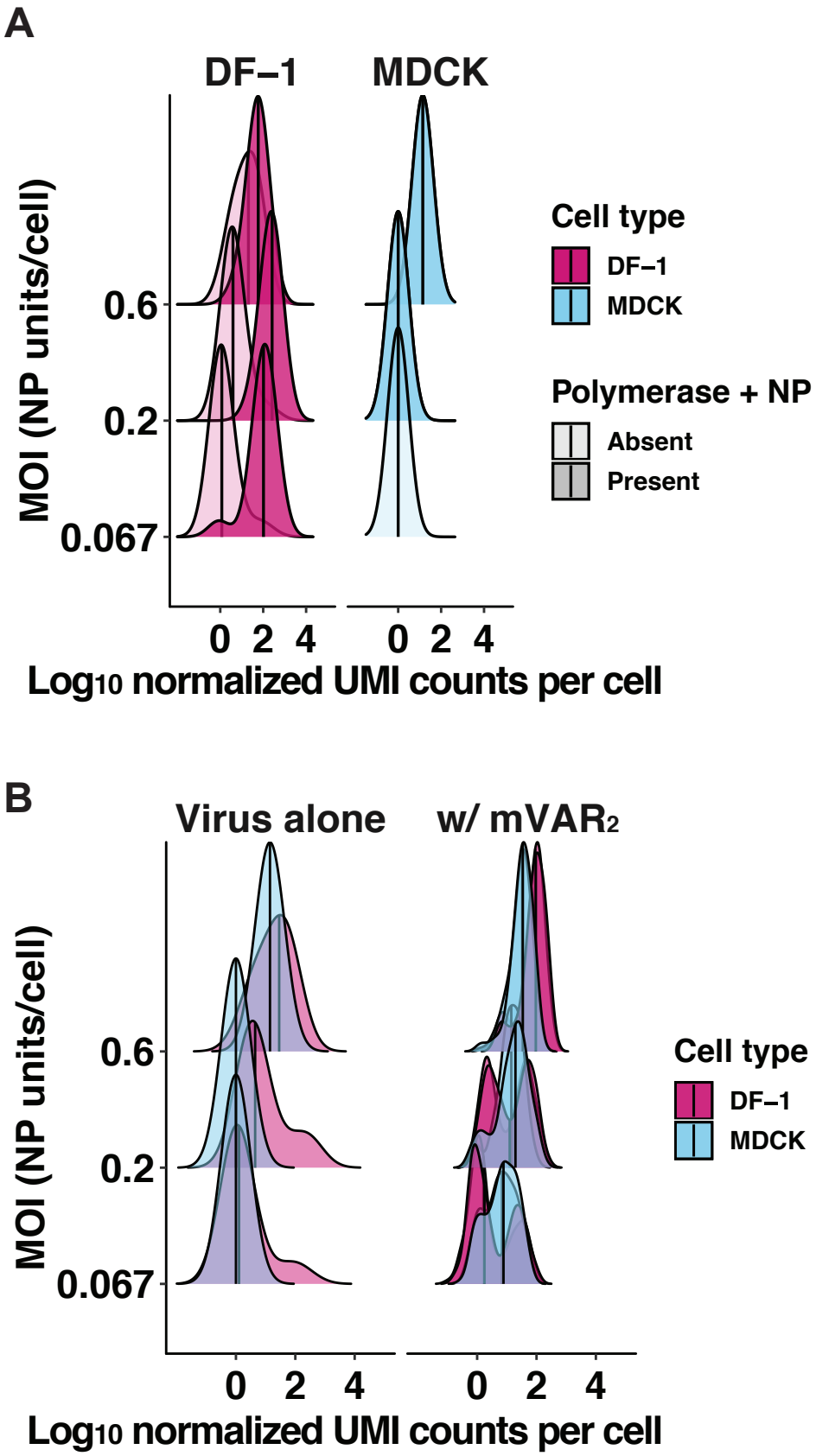

Figure 7