Deep mutational scanning of influenza A virus NEP reveals pleiotropic mutations in its N-terminal domain

Qi Wen Teo[1,2,*], Yiquan Wang[1,*], Huibin Lv[1,2,*], Kevin J. Mao[1], Timothy J.C. Tan[3], Yang Wei Huan[1], Joel Rivera-Cardona[4], Evan K. Shao[1], Danbi Choi[1], Zahra Tavakoli Dargani[1], Christopher B. Brooke[1,2,4], Nicholas C. Wu[1,2,3,5,§]

[1] Department of Biochemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[2] Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[3] Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[4] Department of Microbiology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[5] Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

* These authors contributed equally to this work

[§] To whom correspondence may be addressed. Email: nicwu@illinois.edu (N.C.W.)

ABSTRACT

The influenza A virus nuclear export protein (NEP) is a multifunctional protein that is essential for the viral life cycle and has very high sequence conservation. However, since the open reading frame of NEP largely overlaps with that of another influenza viral protein, non-structural protein 1, it is difficult to infer the functional constraints of NEP based on sequence conservation analysis. Besides, the N-terminal of NEP is structurally disordered, which further complicates the understanding of its function. Here, we systematically measured the replication fitness effects of >1,800 mutations of NEP. Our results show that the N-terminal domain has high mutational tolerance. Additional experiments demonstrate that N-terminal domain mutations pleiotropically affect viral transcription and replication dynamics, host cellular responses, and mammalian adaptation of avian influenza virus. Overall, our study not only advances the functional understanding of NEP, but also provides insights into its evolutionary constraints.

INTRODUCTION

Influenza A virus (IAV) belongs to the *Orthomyxoviridae* family and is characterized as an enveloped, negative-sense single-stranded, segmented RNA virus. Its genome is comprised of eight viral segments encoding at least 12 proteins[1]. Segment eight, which is the shortest segment, is known as the non-structural (NS) segment. NS segment encodes two proteins via splicing, NS protein 1 (NS1) and nuclear export protein (NEP, also known as NS2)[2]. NS1 is one of the earliest proteins expressed during virus infection and has been well characterized as a potent type I interferon (IFN) antagonist[3]. On the other hand, NEP serves a nuclear export function and is expressed at a later stage during infection as a less abundant spliced product of the NS segment[4]. While IAV with NS1 deletion is replication-competent albeit attenuated[5], IAV lacking NEP results in a replication-incompetent virus[6].

NEP has a protease-sensitive N-terminal domain (amino acids 1-53), which harbors two nuclear export signals (NES)[7, 8], and a protease-resistant C-terminal domain (amino acids 54-121)[9], which interacts with M1. A major function of NEP is to mediate the nuclear export of influenza viral ribonucleoproteins (vRNPs) for viral packaging[10]. Each influenza vRNP comprises viral RNA (vRNA) binding to three viral polymerase subunits PB2, PB1, and PA, as well as nucleoprotein (NP). NEP orchestrates the nuclear export of vRNPs through its interaction with cellular β-importin protein CRM1[7, 8] and influenza matrix protein (M1), which in turn binds to the vRNPs[7, 9]. The vRNP complex is responsible for both transcription and replication of the vRNA. Transcription occurs inside the nucleus where the vRNA is transcribed into capped and polyadenylated messenger RNA (mRNA) by a primer-dependent mechanism[11, 12]. For viral genome replication, positive-sense complementary RNA (cRNA) is first produced using vRNA as a template and subsequently serves as a template to produce more vRNA[11, 13]. The synthesis of these three viral RNA species, namely mRNA, cRNA, and vRNA, exhibits distinct dynamics during infection[14, 15]. The timing of these dynamics is important for optimal production

of infectious virions[16, 17]. Several studies have suggested that NEP is a polymerase-associated cofactor that acts as a 'molecular timer' to regulate the switch from transcription to replication, which is attributed to the gradual accumulation of NEP[17, 18, 19].

Viral polymerase activity represents a major barrier for avian IAV adaptation to mammalian hosts. Several mammalian adaptive mutations in the polymerase subunits of avian IAV have been identified, as exemplified by E627K in PB2[20]. However, NEP has received increasing attention as a crucial determinant of IAV tropism, facilitating avian IAV to overcome the replication block in mammalian cells[21, 22, 23]. Notably, avian H5N1 polymerase is particularly susceptible to the polymerase-enhancing ability of NEP that harbor mammalian adaptive mutations (e.g., M16I, Y41C, and E75G)[22]. These mammalian adaptive mutations in NEP significantly augment cRNA and vRNA synthesis, as well as mRNA transcription[22].

While residues 63-116 in the C-terminal domain of NEP form a helical hairpin[9], the rest of the protein is structurally disordered[9, 24]. This poses a challenge to understand its sequence-function relationship, especially for the N-terminal domain. In this study, the replication fitness effects of ~2,000 mutations of NEP were measured using deep mutational scanning. Our findings revealed that the N-terminal domain of NEP displays a greater mutational tolerance than its C-terminal domain. Furthermore, we showed that multiple mutations in the N-terminal domain simultaneously affect the transcription-to-replication switch, innate immune response, cellular apoptosis, as well as facilitate mammalian adaptation of avian IAV.

RESULTS

Deep mutational scanning of H1N1 A/WSN/1933 NEP

To perform deep mutational scanning of NEP, we utilized the eight-plasmid reverse genetics system of influenza H1N1 A/WSN/1933 (WSN) virus with NS1 and NEP separated by a 2A

autoproteolytic cleavage site (**Figure S1A**)[25, 26]. Here, the open reading frames of NS1 and NEP did not overlap, as opposed to NEP being produced as a spliced product of NS1 in the unengineered influenza virus. This virus, herein referred as wild-type NS-split ($WT_{NS\text{-}split}$), enabled us to construct a saturation mutagenesis library spanning residues 2-113 of NEP without mutating NS1. Over 98% of the clones in the mutant library contained zero to one amino acid mutation (**Figure S1B**). The virus mutant library was then rescued and passaged once for 24 hours in MDCK-SIAT1 cells. Using next-generation sequencing, the frequencies of individual mutations in the plasmid mutant library and the post-passaged mutant library were measured. Subsequently, the fitness value of each mutation was computed as the normalized $\log_{10}$ enrichment ratio such that the fitness value of $WT_{NS\text{-}split}$ was 0, whereas beneficial and deleterious mutations would have positive and negative fitness values, respectively. Our deep mutational scanning experiment measured the replication fitness effects of 1,895 (89%) out of 2,128 all possible amino acid mutations across the 112 residues of interest (**Figure 1 and Table S1**). A Pearson correlation of 0.77 was obtained between two biological replicates (**Figure S1C**), demonstrating a reasonable reproducibility of our deep mutational scanning result.

To further validate our deep mutational scanning result, we constructed seven amino acid mutations individually, including three that had low fitness values (M52F, H56W, and Q59T), and four that had high fitness values (I32T, T33L, S37Y, and D43V). Except residue Ile32, the other residues are evolutionarily highly conserved (**Figure S2**). Subsequently, a virus rescue experiment was performed to assess the replication fitness of these mutations (**Figure 2A**). Consistent with our deep mutational scanning result, mutations I32T, T33L, S37Y, and D43V were rescued to a titer similar to or higher than $WT_{NS\text{-}split}$, whereas M52F and Q59T did not produce any detectable infectious particles. Nevertheless, in contrast to the deep mutational scanning result, H56W was rescued to a titer similar to $WT_{NS\text{-}split}$. Together, this validation experiment indicates that our deep mutational scanning result contains some measurement

errors but is mostly reliable. Of note, as described in a previous study[17], $WT_{NS\text{-}split}$ had a two-log decrease in virus titer compared to the parental $WT_{non\text{-}NS\text{-}split}$ (**Figure 2A**).

Enrichment of mutations with high replication fitness at residue 43

Our deep mutational scanning experiment revealed that the C-terminal domain (average mutational fitness value = -0.36) exhibited lower mutational tolerance than the N-terminal domain (average mutational fitness value = 0, **Figure 1**). This may be partly due to the presence of structural constraints in the C-terminal domain but not the N-terminal domain[9, 24]. The C-terminal domain comprises two α-helices C1 (amino acids 64-85) and C2 (amino acids 94-115), connected by a short interhelical turn (amino acids 86-93)[9]. These helices, largely hydrophobic, form extensive contacts throughout their entire length[9]. The C2 displayed lower mutational tolerance than C1, substantiating previous studies showing that C2 is functionally important in modulating influenza polymerase activity[22, 27] and regulating the intramolecular interaction with NEP N-terminal domain[23, 28, 29]. Additionally, C2-encoding region is critical for packaging of NS vRNA into the viral particle[30]. Our result also showed that the key M1-binding residue Trp78 in C1 had a low mutational tolerance except for aromatic amino acid mutations (His, Phe, and Tyr), indicating the importance of aromaticity in NEP-M1 interaction.

The N-terminal domain of NEP is featured by two NES (NES1 and NES2)[8, 10]. A previous study has demonstrated that while the NES1 sequence can tolerate some mutations without abolishing its function, certain hydrophobic residues (Met14, Met16, Met19 and Leu21) are crucial[31]. Our data corroborated this finding, revealing that although most mutations within the NES were viable, mutations at hydrophobic residue Leu21 in NES1 were predominantly deleterious (**Figure 1**). In contrast to Leu21, most mutations at residue Asp43, which did not reside in any region with known functions, exhibited high replication fitness in our deep mutational scanning result (**Figure 1**). To experimentally validate this finding, all 19 amino acid

mutations were individually introduced into residue 43 of the $WT_{NS\text{-}split}$ and analyzed by a virus rescue experiment. Consistent with our deep mutational scanning result, our virus rescue experiment showed that except Glu, which is chemically similar to Asp, all other amino acids mutations increased the titer by 0.5-3 log (**Figure 2B**). In fact, the titer of individual mutants in our virus rescue experiments had a Pearson correlation of 0.78 with the deep mutational scanning result (**Figure S1D**). Collectively, while the N-terminal domain had a high mutational tolerance, certain mutations in the N-terminal domain could dramatically alter virus fitness, highlighting its functional significance.

N-terminal domain mutations modulate replication dynamics

To delve deeper into the functional relevance of the N-terminal domain in the IAV life cycle, we further characterized those five replication-competent mutations in our first virus rescue experiment (**Figure 2A**). These included I32T, T33L, S37Y, D43V, which resided within the N-terminal domain, and H56W, which was in the structurally disordered region downstream of the N-terminal domain. Specifically, we were interested in understanding how mutations I32T and D43V increased the virus titer (**Figure 2A**). Since NEP has been shown to modulate the production of defective interfering particles (DIPs)[32], we first investigated the DIP production of different mutants by next-generation sequencing. However, our data indicated that the DIPs could not explain the differential virus production among different mutants (**Figure S3A**). For example, although the titer of $I32T_{NS\text{-}split}$ and $D43V_{NS\text{-}split}$ was a log higher than $WT_{NS\text{-}split}$ in our virus rescue experiment (**Figure 2A**), their DIP production profiles were similar.

A major function of NEP is to regulate switching of viral transcription to replication[16, 17, 18, 33]. Therefore, we hypothesized that the difference in virus replication fitness among these mutants was due to their effects on IAV transcription and replication dynamics. To test our hypothesis, we employed quantitative reverse transcription-PCR (RT-qPCR). Although RT-qPCR is routinely

used to probe IAV kinetic profile, it has several drawbacks, including high background and low specificity[34]. Consequently, to complement RT-qPCR, we performed the influenza virus enumerator of RNA transcripts (InVERT) analysis, which is an RNA sequencing (RNA-seq) approach that enables quantification of all three RNA species produced by IAV[14]. Both RT-qPCR and InVERT analyses revealed that the transcription-to-replication switch for $WT_{NS\text{-}split}$ occurred earlier than the parental $WT_{non\text{-}NS\text{-}split}$. Moreover, $I32T_{NS\text{-}split}$ and $D43V_{NS\text{-}split}$, which had higher replication fitness than $WT_{NS\text{-}split}$ (**Figure 2A**), displayed transcription-to-replication switch at a later stage of infection, more closely resembling the parental $WT_{non\text{-}NS\text{-}split}$ (**Figure 3A and S3B**). Conversely, $T33L_{NS\text{-}split}$, which had similar replication fitness as $WT_{NS\text{-}split}$ in our rescue experiment (**Figure 2A**), had the switch at an earlier stage of infection (**Figure 3A and S3B**).

Since NEP facilitates nuclear export of vRNPs[7], we also performed RNA fluorescence *in situ* hybridization (FISH) to examine if the nuclear export of vRNPs was affected by these mutants. In this experiment, probes that were specific to PB2 vRNA were used. Notably, cells that were infected with $WT_{NS\text{-}split}$ and $T33L_{NS\text{-}split}$ showed bright perinuclear foci of PB2 vRNA (**Figure 3B and 3C**). This was in stark contrast with the vRNA distribution in cells that were infected with $WSN_{non\text{-}NS\text{-}split}$ and $D43V_{NS\text{-}split}$, where the vRNA was predominantly localized within the nucleus (**Figure 3B and 3C**). These results indicated that the D43V substitution also affected the vRNPs' nuclear export, despite not being positioned within the NES. Additionally, the timing disparity in the transcription-to-replication switch and the altered vRNPs nuclear export between $WT_{NS\text{-}split}$ and $D43V_{NS\text{-}split}$ suggested that the increase in replication fitness of D43V may be attributed to altered replication kinetics.

Considering the importance of the NS1 to NEP expression ratio in coordinating the timing of IAV infection[17], we postulated that the variation of transcription-to-replication switch among mutants correlated with NS expression. Consistently, the parental $WT_{non\text{-}NS\text{-}split}$ showed a higher NS1 and

lower NEP expression compared to $WT_{NS\text{-}split}$ (**Figure 3D**). NS1 and NEP expression patterns of mutants with higher replication fitness, namely $I32T_{NS\text{-}split}$ and $D43V_{NS\text{-}split}$, resembled those of $WT_{non\text{-}NS\text{-}split}$. By contrast, mutants with lower replication fitness, namely $T33L_{NS\text{-}split}$, $S37Y_{NS\text{-}split}$, and $H56W_{NS\text{-}split}$, showed higher NEP and lower NS1 expressions (**Figure 3D**). These findings suggested that NEP mutations I32T and D43V compensated the fitness defect of $WT_{NS\text{-}split}$ by suppressing the NEP expression level to that of $WT_{non\text{-}NS\text{-}split}$, which may in turn lead to transcription-to-replication switch dynamics akin to that of $WT_{non\text{-}NS\text{-}split}$.

Impact of mutations in the N-terminal domain on cellular response regulation

Next, we examined the effects of these mutants on multicycle replication kinetics in both MDCK-SIAT1 and A549 cells. Virus production of $I32T_{NS\text{-}split}$ and $D43V_{NS\text{-}split}$ was approximately one log higher than $WT_{NS\text{-}split}$ at 48 hours post-infection (hpi) in both cell lines (**Figure 4A and S4A**). Notably, although virus production of $T33L_{NS\text{-}split}$ was comparable to $WT_{NS\text{-}split}$ in MDCK-SIAT1 cell, it was significantly lower than $WT_{NS\text{-}split}$ in A549 cells (**Figure 4A and S4A**). The disparity in virus production of $T33L_{NS\text{-}split}$ between A549 and MDCK-SIAT1 may be due to differential interferon (IFN) induction in these cell types[35]. These results prompted us to hypothesize that these mutants may affect cellular response during infection, thereby contributing to their different growth kinetics.

A recent study has indicated that NEP promotes innate immunity evasion by interacting with interferon regulatory factor (IRF7) through the N-terminal domain[36]. To ascertain the impact of the NEP mutants on innate immunity, we infected $A549^{Dual}$ cells which allow simultaneously monitoring of IRF activity and NF-κB induction[37]. Our data demonstrated that $I32T_{NS\text{-}split}$ and $D43V_{NS\text{-}split}$ had significantly lower IRF and NF-κB induction compared to the $WT_{NS\text{-}split}$ at 24 and 48 hpi (**Figure 4B and S4B**). This observation helps explain the higher virus production in

I32T$_{NS\text{-}split}$ and D43V$_{NS\text{-}split}$ compared to the WT$_{NS\text{-}split}$ at 48 hpi (**Figure 4A and S4A**). NEP has also been shown to suppress ISRE activation[36]. As a result, we further tested if these NEP mutants affect ISRE activity by transiently transfecting HEK293T cells with NEP expression constructs, an ISRE-firefly luciferase reporter plasmid, and a constitutively expressed Renilla luciferase plasmid for normalization. Consistently, we showed that NEP suppressed ISRE activity in response to recombinant IFN-α (**Figure 4C**). This inhibition can be observed across all NEP mutants with varying protein expression levels (**Figure 4C and S4C**). Our result also indicated that T33L mutation weakened the ability of NEP to antagonize IFN response despite comparable, if not higher, protein expression level with other mutants (**Figure 4C**). This helps explain the lower virus production observed in T33L$_{NS\text{-}split}$ compared to that of WT$_{NS\text{-}split}$ in A549 cells (**Figure 4A**).

When the innate immune responses fail to effectively control IAV replication, cells may activate a secondary antiviral response via programmed death, known as apoptosis[38]. However, IAV strategically inhibits early induction of apoptosis to prolong replication and induces apoptosis at later stages of infection to enhance viral spread[39, 40, 41, 42]. Given the difference in innate immune response among mutants and unexpected extensive cell death observed in some mutants (WT$_{NS\text{-}split}$, T33L$_{NS\text{-}split}$, S37Y$_{NS\text{-}split}$, and H56W$_{NS\text{-}split}$) (**Figure S4D**), we next examined the apoptosis pathway in infected A549 cells. Our results revealed that while the parental WT$_{non\text{-}NS\text{-}split}$, I32T$_{NS\text{-}split}$ and D43V$_{NS\text{-}split}$ induced apoptosis at 9 hpi, WT$_{NS\text{-}split}$, T33L$_{NS\text{-}split}$, S37Y$_{NS\text{-}split}$, and H56W$_{NS\text{-}split}$ infection resulted in biphasic induction of apoptosis, occurring as early as 6 hpi and again at 16 hpi **(Figure 4D**). Therefore, the differential viral RNA transcription-to-replication switch among these mutants may not solely account for their difference in replication fitness due to the intricate interplay between virus and cellular responses. Taken together, our findings suggest that the N-terminal domain of NEP modulates virus production through regulating multiple cellular responses.

240

241 **NEP mutations increase the polymerase activity of avian IAV in human cells**

242 Previous studies have shown that mutations in the C-terminal domain of NEP can modulate the
243 polymerase activity[23, 43]. In this study, we sought to explore whether the N-terminal domain also
244 influenced polymerase activity. We first performed an experiment by testing the effects of our
245 NEP mutants on the WSN polymerase, using a polymerase reconstitution assay in HEK293T
246 cells. Nevertheless, no discernible differences in WSN polymerase activity were observed
247 across all tested NEP mutants, irrespective of the amount of NEP present (**Figure 5A**).

248

249 However, given that there is growing evidence indicating that mammalian-adaptative mutations
250 in NEP enhance avian IAV replication in cultured human cells[21, 22, 23, 44, 45], we aimed to further
251 investigate the effect of these NEP mutants on the avian IAV polymerase activity in human cells.
252 We individually introduced each of the mutations I32T, D43V, and H56W into NEP from a low
253 pathogenic avian influenza (LPAI) strain, A/Quail/Hong Kong/G1/97 (H9N2), and a highly
254 pathogenic avian influenza (HPAI) strain, A/Vietnam/1203/2004 (H5N1). Subsequently, we
255 assessed their polymerase activity in HEK293T cells. Our data revealed that all three NEP
256 mutants partially increased the polymerase activity of the H9N2 and H5N1 strains (**Figure 5B**
257 **and 5C**). Surprisingly, H56W mutant also enhanced the avian polymerase activity (**Figure 5B**
258 **and 5C**), despite having very different phenotypes compared to I32T and D43V mutants in
259 terms of protein expression level (**Figure 3D and S3C**), IRF induction (**Figure 4B**), and
260 apoptosis induction (**Figure 4D**). In sum, these findings suggest that I32T, D43V and H56W
261 mutations in the NEP of H9N2 and H5N1 can potentially facilitate mammalian adaptation of
262 avian IAV.

263

264 **DISCUSSION**

For years, NEP has been perceived as a non-structural protein when purified from virions[46, 47]. It was not until 1988 that the nuclear export function of NEP was unveiled[7], which represented a pivotal discovery in influenza virus biology. As our understanding of NEP has evolved, its multifaceted roles have emerged as focal points of investigation. However, more than 25 years since the discovery of its nuclear export function, the sequence-function relationships of NEP remain largely obscure. By systematically defining the mutational fitness landscape of influenza NEP, this study advances our understanding of its functional constraints.

A highlight of the study is that mutations at N-terminal domain of NEP display pleiotropic effects including viral RNA synthesis dynamics, innate immune response modulation, apoptosis induction, and mammalian-adaptation of avian IAV. We speculate that the pleiotropic phenotypes of the N-terminal domain mutations stem from alterations in the NEP:NS1 expression ratio, which subsequently perturbs viral RNA synthesis dynamics and cellular responses. Nonetheless, the modulation of apoptosis by NEP mutations was not expected. While NS1 is known to be an important regulator of host apoptotic cell death[39, 41], the role of NEP in apoptosis is less clear. Although our study shows that the NEP mutant viruses induce monophasic or biphasic apoptosis, it is possible that the NS1 expression level is the primary leading cause of this phenotype. Future investigation is required to dissect the underlying mechanisms.

One interesting aspect of overlapping open reading frames is how nucleotide usage is balanced to according to the functions of each open reading frame[48]. A previous study has found that the evolution of the overlapping regions between Tat and Rev of HIV-1 is constrained by the function of either but not both of the proteins[48]. In the case of NS protein, majority of the N-terminal domain of NEP overlaps with the C-terminal effector domain of NS1, which predominantly suppresses the host response and is associated with the pathogenicity of the

291 virus[49, 50]. Given the high mutational tolerance in the disordered N-terminal of NEP and the role
292 of NS1 as a key immunomodulatory factor, it is likely that NEP accommodates the nucleotide
293 preferences of NS1 due to functional constrains imposed by NS1. In other words, the sequence
294 conservation of the N-terminal domain is largely due to the overlapping open reading frames
295 with NS1, rather than its own functional constraints. Of note, overlapping open reading frames
296 of NS proteins necessitates the use of the 'NS-split' system in the present study. However, this
297 hampers our understanding of how mutations within this domain would impact virus replication
298 kinetics as well as cellular responses in natural infection.
299
300 Most reported mammalian-adaptive mutations in avian IAV NEP occur in the N-terminal
301 domain[22, 23], although previous studies have highlighted the importance of the three amino acid
302 at the C-terminal end of NEP for its polymerase-enhancing function[23, 43]. Additionally, mutation
303 T48N in the N-terminal domain of NEP from pandemic H1N1/2009-like recombinant virus
304 (rH1N1pdm) has been found to enhance viral replication in guinea pigs[51], further consolidating
305 our hypothesis that N-terminal domain of NEP is an important tropism determinant. One notable
306 observation in this study is that all tested mutations in the N-terminal domain of NEP were found
307 to enhance the polymerase activity of avian IAV but not WSN in mammalian cells. A previous
308 study has suggested that NEP has two distinct conformations, "open" and "closed", and that
309 mammalian adaptive mutations in NEP shift the equilibrium towards the "open" conformation[23].
310 In our study, we used the lab-adapted A/WSN/33 strain, which is known for robust replicates in
311 mammalian cell culture. Therefore, WSN NEP may already favor the "open" conformation
312 without additional mammalian adaptive mutations, although this speculation requires detailed
313 investigations in the future. Additionally, whether mammalian-adaptive mutations in NEP
314 influence its interaction with the vRNP, thus leading to enhanced polymerase activity, remains
315 an unanswered question.
316

ACKNOWLEDGEMENTS

We thank the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign for assistance with next-generation sequencing and RNA sequencing. We are grateful to the Institute of Genomic Biology Core Facility at the University of Illinois at Urbana-Champaign for providing access to the LSM880 microscope system. This work was supported by the Searle Scholars Program (N.C.W.) and National Institutes of Health (NIH) R01 AI165475 (N.C.W.) and R01 AI139246 (C.B.B.).

AUTHOR CONTRIBUTIONS

Q.W.T., Y.W., H.L., and N.C.W. conceived and designed the study. Y.W. constructed the NEP deep mutational scanning library and prepared the sequencing libraries. Q.W.T. prepared the samples for RNA sequencing. Q.W.T., Y.W., and N.C.W. performed data analysis. Q.W.T., H.L., K.J.M., Y.W.H., T.J.C.T., J.R.-C., E.K.S., D.C., Z.T.D., and C.B.B. assisted with experiments. Q.W.T., Y.W., H.L., and N.C.W. wrote the paper and all authors reviewed and/or edited the paper.

DECLARATION OF INTERESTS

N.C.W. consults for HeliXon. The authors declare no other competing interests.

FIGURE LEGENDS

Figure 1. Deep mutational scanning of WSN NEP. The replication fitness effects of individual mutations in the WSN NEP were measured by deep mutational scanning and are shown as a heatmap. Wild-type (WT) amino acids are indicated by a black circle. Mutations in gray were excluded in our data analysis due to low input count. N-terminal domain, C-terminal domain, nuclear export signal (NES), and the M1-binding site are annotated by colored bars. Two α-

342 helices C1 and C2 are annotated by pink and blue bars, respectively. Stop codon is annotated
343 as *.
344
345 **Figure 2. Virus rescue experiment of mutations at residue Asp43. (A-B)** Mutations were
346 individually introduced into WSN NEP. Their effects on replication fitness were examined by a
347 virus rescue experiment. Virus titer was measured by $TCID_{50}$ assay. **(A)** Replication fitness
348 effects of selected mutations at various residues. Each bar represents the mean of two
349 independent biological replicates. Statistical significance was analyzed by two-tailed Student's *t*-
350 test: $*P < 0.05$. **(B)** Replication fitness effects of mutations at residue Asp43. Each bar
351 represents the mean of three independent biological replicates. Difference in virus titer between
352 $WT_{NS\text{-}split}$ and mutations at residue 43 is statistically significant, except for D43E. Statistical
353 significance was analyzed by two-tailed Student's *t*-test: ns: not significant. Different batches of
354 cells were used for the virus rescue experiment in (A) and (B).
355
356 **Figure 3. Effects of D43V and T33L mutations on transcription and replication. (A-D)**
357 MDCK-SIAT1 cells were infected with the indicated mutants at an MOI of 5. **(A)** Kinetics of
358 different IAV RNA species during infection are shown. RNA from infected cells were harvested
359 at the indicated timepoints and analyzed by RT-qPCR. The specificities of the amplified
360 products were confirmed by the melting curve analysis. Data are shown as means of three
361 independent biological replicates. Transcription-to-replication switch was indicated with a
362 vertical dotted line. The switch was defined as the point where mRNA expression intersected
363 with the increasing cRNA expression or reached its plateau level. Blue: vRNA; Orange: cRNA;
364 Green: mRNA. **(B)** Micrographs of FISH staining of the indicated mutants are shown. Infected
365 cells were fixed at 9 hpi, and then stained for PB2 vRNA (orange) and nucleus (cyan). Scale bar:
366 50 μm. **(C)** Ratio of mean intensity of PB2 vRNA in the nucleus to that of PB2 vRNA in the
367 cytoplasm for micrographs in panel B. Each bar represents the mean ± standard error with

individual data points shown (n = 46 for $WT_{non\text{-}NS\text{-}split}$; n = 29 for $WT_{NS\text{-}split}$; n = 39 for $T33L_{NS\text{-}split}$; n = 91 for $D43V_{NS\text{-}split}$). Two outlier data points were removed for cells infected with $WT_{non\text{-}NS\text{-}split}$, $T33L_{NS\text{-}split}$, and $D43V_{NS\text{-}split}$, respectively. Statistical significance was analyzed by two-tailed Student's *t*-test: ns: not significant, ***P* < 0.01, and *****P* < 0.0001. **(D)** Western blot analysis of NP, NS1, and NEP protein expression at 16 hpi. GAPDH was used as loading control.

Figure 4. Impacts of NEP mutations on cellular responses. (A) A549 cells were infected at an MOI of 0.01 with the indicated mutants. Infectious particles in the supernatant were harvested at the indicated timepoints and quantified via the $TCID_{50}$ assay. Each bar represents the mean of three independent biological replicates. Statistical significance was analyzed by two-way ANOVA: **P* < 0.05, ***P* < 0.01, and *****P* < 0.0001. **(B)** Activation of IRF activity was measured in mock- and IAV-infected A549 cells expressing dual reporters for NF-κB and IRF activities. Each bar represents the mean of three independent biological replicates. Statistical significance was analyzed by Student's *t*-test: ****P* < 0.001 and *****P* < 0.0001. **(C)** HEK293T cells were transiently transfected with plasmids expressing the indicated NEP mutants, a plasmid encoding an ISRE-firefly luciferase reporter, and a plasmid expressing Renilla luciferase. At 24 h post-transfection, cells were treated with IFN-I (2000 U/mL) for 16 h prior to measuring luciferase activities. Each bar represents the mean of three independent biological replicates. Statistical significance was analyzed by two-tailed Student's *t*-test: ****P* < 0.001. **(D)** A549 cells were infected with the indicated mutants at an MOI of 5. Cells were harvested at the indicated timepoint, and the number of cells undergoing apoptosis were measured using CellEvent Caspase-3/7 Green Flow cytometry kit. Each line represents the mean of three independent biological replicates. Cleaved Caspase-3/7: c-Casp3/7.

Figure 5. Polymerase activity of avian influenza in cultured human cells. Influenza polymerase reconstitution assay in HEK293T cells for **(A)** WSN, **(B)** A/Quail/Hong Kong/G1/97

(H9N2) and **(C)** A/Vietnam/1203/2004 (H5N1). HEK293T cells were transiently transfected with plasmids encoding the RNP complexes (PB1, PB2, PA, and NP) and vector or indicated amount of plasmids encoding NEP mutants, together with human polymerase I-driven vRNA and Renilla-expressing luciferase reporters. Renilla activity was used to normalize variation in transfection efficiency. Measurements were taken at 16 h post-transfection. Each bar represents the mean of three independent biological replicates. Statistical significance was analyzed by Student's *t*-test: **P* < 0.05, ***P* < 0.01, ****P* < 0.001 and *****P* < 0.0001.

METHODS

Cells and viruses

Madin-Darby canine kidney cells with stable expression of human α-2,6-sialyltransferase (MDCK-SIAT1) (Sigma, catalog #: 05071502-1VL), human lung adenocarcinoma epithelial A549-Dual cells (InvivoGen, catalog #: a549-nfis) and human embryonic kidney HEK293T (ATCC, catalog #: CRL-3216) cells were maintained in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10% fetal bovine serum (FBS; Gibco), 100 U/mL Penicillin-Streptomycin. Cells were cultured in humidified incubators at 37 °C with 5% CO_2. Recombinant influenza A/WSN/1933 (H1N1/WSN) virus was generated using the pHW2000 eight-plasmid reverse genetics system[26]. Viral genotypes were confirmed by sequencing vial genomes.

Mutant library construction

The mutant library was constructed based on the pHW2000 eight-plasmid reverse genetics system for influenza virus. The linearized vector and a library of mutant WSN NEP inserts were generated through PCR. The linearized vector was generated using 5'- CGA CCG TCT CTG GGG CCG GGA GGT CGC GTC ACC GAC-3' and 5'- CGA CCG TCT CTT CTC CTC GAC GTC CCC GGC TTG CTT-3' as primers. For insert generation, two batches of PCRs were performed, followed by overlapping PCRs. The first batch of PCRs consisted of 15 reactions,

each with an equal molar mix of eight forward primers, as the forward primer and a universal reverse primer 5'- CGA CCG TCT CTC CCC GGG GGA GGT ATA TCT TT-3'. The forward primers for the first batch of PCRs are listed in Table S2. These forward primers were named as cassetteX_N, in which X represents the cassette ID and N represents the primer number. Forward primers with the same cassette ID were mixed at equal molar ratios and used in the same single PCR. The second batch of PCRs consisted of another 15 reactions, each with a universal forward primer 5'- CGT TAC CCG GCC AAT GCA CGT TTA CGC CAC AAA TTT CTC TCT CCT CAA GCA AGC CGG GGA CGT CTC GGA GAA TCC CGG GCC C-3'. and unique reverse primers as listed in Table S2. Subsequently, 15 overlapping PCRs were performed using the universal forward primer and reverse primer. For each overlapping PCR, the template was a mixture of 10 ng each of the corresponding products from the first and second batches of PCRs. The complete insert was an equal molar mix of the products of these 15 overlapping PCRs. All PCRs were performed using PrimeSTAR Max polymerase (Takara Bio) according to the manufacturer's instructions. PCR products were purified using Monarch DNA Gel Extraction Kit (New England Biolabs). Both the vector and the complete insert were digested by BsmBI (New England Biolabs) and ligated using T4 DNA ligase (New England Biolabs). The ligated product was transformed into MegaX DH10B T1R cells (Thermo Fisher Scientific). At least one million colonies were collected. Plasmid mutant libraries were purified from the bacteria colonies using PureLink HiPure Expi Plasmid Purification Kit (Thermo Fisher Scientific).

Rescuing, passaging and sequencing the mutant library

The virus mutant library was rescued by transfecting a co-culture of HEK293T and MDCK-SIAT1 cells (ratio of 6:1) at 60% confluence in a T75 flask. Transfection was performed using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. At 24 h post-transfection, cells were washed twice with PBS and cell culture

446 medium was replaced with OPTI-MEM medium supplemented with 0.8 μg/ml TPCK-trypsin. The
447 virus mutant library was harvested at 72 h post-transfection, titered by $TCID_{50}$ assay using
448 MDCK-SIAT1 cells then stored at –80°C until use. To passage the virus mutant libraries,
449 MDCK-SIAT1 cells in a T75 flask were washed twice with PBS and then infected with a
450 multiplicity of infection (MOI) of 0.02 in OPTI-MEM medium supplemented with 0.8 μg/ml TPCK-
451 trypsin. At 2 h post-infection, infected cells were washed twice with PBS and fresh OPTI-MEM
452 medium supplemented with 0.8 μg/ml TPCK-trypsin was added to the cells. At 24 h post-
453 infection, supernatant was harvested. Each replicate was transfected and passaged
454 independently.
455
456 Viral RNA of the post-passaged mutant library was extracted using QIAamp Viral RNA Mini Kit
457 (QIAGEN). The extracted RNA was then reverse transcribed to cDNA using Superscript III
458 reverse transcriptase (Thermo Fisher Scientific). To add part of the adapter sequence required
459 for Illumina sequencing, the plasmid mutant library and the cDNA from the post-infection viral
460 mutant library were amplified by PCR using primers: 5'- ACT CTT TCC CTA CAC GAC GCT
461 CTT CCG ATC TNN NNN NNN GAC GTC GAG GAG AAT CCC GGG -3' and 5'-ACT GGA
462 GTT CAG ACG TGT GCT CTT CCG ATC TNN NNG TAG AAA CAA GGG TGT TTT TTA TTA-
463 3'. A total of 12 N were included in each amplicon product as unique molecular identifiers (UMIs)
464 to distinguish true mutations from sequencing errors[52, 53, 54]. For each sample, the six amplicon
465 PCR products were mixed at equal molar ratio. Subsequently, 1 million copies of mixed
466 amplicon PCR products were used as template for a second PCR to add the rest of the adapter
467 sequence and index to the amplicon using primers: 5'-AAT GAT ACG GCG ACC ACC GAG
468 ATC TAC ACX XXX XXX XAC ACT CTT TCC CTA CAC GAC GCT-3', and 5'-CAA GCA GAA
469 GAC GGC ATA CGA GAT XXX XXX XXG TGA CTG GAG TTC AGA CGT GTG CT-3'.
470 Positions annotated by an "X" represented the nucleotides for the index sequence. The final

PCR products were purified by PureLink PCR purification kit (Thermo Fisher Scientific) and submitted for next-generation sequencing using Illumina MiSeq PE250.

Analysis of sequencing data for deep mutational scanning

Sequencing data was obtained in FASTQ format and analyzed using a custom snakemake pipeline[55]. Firstly, paired-end reads with the same Unique Molecular Identifiers (UMIs) were merged using a Python script. A consensus sequence was created for UMIs with a minimum of two identical sequences. Groups were retained if at least 70% of the sequences agreed on the consensus; otherwise, the group was discarded. Next, primer sequences were removed using cutadapt[56], followed by the merging of forward and reverse consensus reads using FLASH[57]. The resulting merged consensus reads were processed using the SeqIO module in BioPython[58], translated into amino acid sequences, and filtered based on their sequence length matching the reference amplicon. Subsequently, mutations were identified by comparing the amino acid sequences to the reference, and merged consensus reads with more than one amino acid mutation were excluded from downstream analysis. The frequency of mutant i in individual samples was normalized as follows:

$$freq_{sample}(i) = \frac{read\ count_{sample}(i) + 1}{\sum_{i \in sample}(read\ count_{sample}(i) + 1)}$$

For each mutant i, the replication fitness was calculated as follows:

$$f_i = \log_{10} \frac{freq_{output}(i)/freq_{input}(i)}{freq_{output}(wt)/freq_{input}(wt)}$$

where the $freq_{output}(i)$ represents the normalized frequency corresponding to variant i in the post-passaging virus mutant library, and the $freq_{input}(i)$ represents the normalized enrichment corresponding to variant i in the plasmid mutant library.

493 **RNA library preparation and sequencing**

494 MDCK-SIAT1 cells were infected with viruses at an MOI of 0.5 and incubated at 4 °C for 30 min
495 to allow the virus to adsorb to the cell surface but not to be endocytosed. Synchronized IAV
496 infection was initiated by shifting the temperature to 37 °C. Secondary infection was blocked
497 using 20 mM NH_4Cl. Cells were harvested at 0, 2, 4, 6, 8, and 12 h post-temperature shift for
498 RNA sequencing to quantify the cellular RNA and viral RNA levels. Uninfected MDCK-SIAT1
499 cells were also collected for a negative control. RNA from the cells were extracted using
500 RNeasy Plus Micro Kit (Qiagen). To assay both positive- and negative-sense RNAs, total RNA
501 was prepped using the Clontech Pico SMARTer stranded total RNA-Seq kit v2, which preserved
502 strand orientation of the RNA sequences by the template-switching reactions. The cDNA library
503 was then subjected to next-generation sequencing using two lanes of Illumina NovaSeq X 10B
504 PE150. At least 20 million paired-end reads were obtained per sample.
505
506 **Quantification of different influenza RNA species**
507 Raw reads underwent adapter trimming using cutadapt[56]. Trimmed reads were then aligned to
508 the reference genome using STAR[59]. The expression levels of all viral genes were initially
509 quantified using htseq-count[60] to count reads that were assigned to different features in the
510 annotation file and calculate fragments per kilobase of transcript per million mapped reads
511 (FPKM) for each feature. Distinction between viral negative-sense RNA (vRNA) and positive-
512 sense RNA (cRNA and mRNA) was achieved by providing an annotation file specifying RNA
513 strandedness during quantification. cRNA and mRNA were differentiated based on the
514 sequence following five adenosines (5A) at the 3' end of the gene. During mRNA transcription,
515 the 5A sequence extends to form the poly(A) tail. In contrast, all cRNAs contain a 16-nt
516 sequence at the 3' end, 13 of which are conserved across segments following the 5A. For each
517 segment in every sample, the pipeline counted sequencing reads that contained the 5A and at
518 least two subsequent nucleotides with complete alignment to the 16-nt consensus sequence as

519 cRNA reads. Reads that contained more than two additional adenosines after the 5A are
520 identified as mRNA reads. The expression values of cRNA and mRNA were determined by
521 multiplying the total read count of positive-sense RNAs.
522
523 **Sequencing analysis of defective interfering particles (DIPs)**
524 Raw sequencing reads were fed into our RNA species pipeline for viral negative-sense RNA
525 (vRNA) quantification. Following alignment to the viral genome, identification and quantification
526 of vRNA containing deletion junctions were performed using a custom Python script. To account
527 for the variation of sequencing read coverage between samples, DIPs percentage was derived
528 by normalizing the deletion junction count to the coverage depth of the corresponding vRNA
529 segments.
530
531 **Reverse transcription and quantitative PCR**
532 MDCK-SIAT1 cells were infected with viruses at an MOI of 0.5 and incubated at 4 °C for 30 min
533 to allow the virus to adsorb to the cell surface but not to be endocytosed. Synchronized IAV
534 infection was initiated by shifting the temperature to 37 °C. Cells were harvested at 0, 2, 4, 6, 8,
535 and 12 h post-temperature shift. Uninfected MDCK-SIAT1 cells were also collected as a
536 negative control. Total cellular RNA was isolated using RNeasy Mini Kit (Qiagen). Reverse
537 transcription (RT) of RNA was performed using ProtoScript II First Strand cDNA Synthesis Kit
538 (New England Biolabs), in accordance with the manufacturer's manual. Oligo-dT primer was
539 used in RT reaction for detection of mRNAs, whereas uni-12- (5'-AGC AAA AGC AGG-3') and
540 uni-13- (5'-AGT AGA AAC AAG G-3') specific primers were used for vRNAs and cRNA,
541 respectively[61, 62]. Quantification PCR (qPCR) mixtures were prepared according to the user
542 manual of iTaq Universal SYBR Green Supermix (Bio-Rad) and reactions were run in a CFX
543 Opus 96 Real-Time PCR Instrument (Bio-Rad). Segment-specific primers were used for the
544 qPCR analysis. WSN PA forward primer: 5'-CTG ACC CAA GAC TTG AAC CAC-3'; WSN PA

545 reverse primer: 5'-AGC ATA TCT CCT ATC TCA AGA ACA-3'; WSN NA forward primer: 5'-ACA
546 ACG GCA TAA TAA CTG AAA CC-3'; WSN NA reverse primer: 5'-CAG GTA CAT TCA GAC
547 TCT TGT GTT-3'. The relative viral copy number quantification was calculated from the slope of
548 a standard curve that was obtained by using serial dilution of the corresponding plasmid as
549 template.
550
551 **Fluorescence *In Situ* Hybridization**
552 One day before the experiment, MDCK-SIAT1 cells were seeded on coverslips in 24-well plates.
553 Cells were infected by viruses at an MOI of 5. Infection was synchronized and secondary
554 infection was blocked using 20 mM NH_4Cl. Coverslips with cells were collected and washed
555 twice with PBS, followed by fixation with 4% formaldehyde for overnight at 4 °C. To visualize
556 viral RNA, RNA fluorescence *in situ* hybridization (RNA-FISH) through hybridization chain
557 reaction (HCR)[63] was then performed according to the manufacturer's instructions with
558 modifications. Cells were washed with PBS thrice, then permeabilized with 0.1% v/v Triton X-
559 100 for 10 min at room temperature. After the permeabilization, cells were washed twice with 2×
560 sodium chloride sodium citrate (SSC) buffer and incubated in HCR hybridization buffer at 37 °C
561 for 30 min with gentle rocking. Subsequently, cells were incubated with hybridization buffer
562 containing 1.2 pmol of the PB2 vRNA probe at 37 °C for overnight with gentle rocking.
563
564 The next day, excess probes were removed by washing cells 4 times with probe wash buffer at
565 37 °C with gentle rocking and were further washed twice with 5× SSC-T buffer (5× SSC with
566 0.1% Tween-20) for 5 min each at room temperature. Cells were then incubated with
567 amplification buffer for 30 min at room temperature with gentle rocking. 18 pmol of B2-Alexa 647
568 amplifier hairpins per sample were snap-cooled by heating at 95 °C for 90 s and cooled to room
569 temperature in the dark for 30 min. Hairpin solution, which was prepared by adding amplifier
570 hairpins to 300 µL amplification buffer per sample, was added to samples and incubated for 1 h

571 at room temperature in the dark with gentle rocking. Excess hairpins were removed by washing
572 samples 5 times with 5× SSC-T buffer. For the first wash, samples were incubated in 5× SSC-T
573 buffer supplemented with 1 µg/mL DAPI (Invitrogen) for 20 min at room temperature with gentle
574 rocking. All subsequent washes were performed for 5 min at room temperature with gentle
575 rocking. After the last wash, samples were mounted on glass slides with ProLong Diamond
576 Antifade Mountant (Invitrogen). Cells were visualized using a LSM880 microscope system
577 (Zeiss).
578
579 **Micrograph Analysis**
580 Micrographs were analyzed using a custom pipeline on CellProfiler v4.2.1[64] (Broad Institute).
581 Using images from the DAPI channel, nuclei were identified using a minimum cross-entropy
582 thresholding method. Then, nuclei were used to propagate and identify cells using the Otsu
583 thresholding method. Nuclei and cells that touch the border of the image were discarded.
584 Cytoplasm was identified by subtracting nuclei from cells. The ratio of the mean intensity of PB2
585 vRNA in the nucleus to that of PB2 vRNA in the cytoplasm was subsequently calculated. Only
586 cells with PB2 vRNA median intensity of at least 0.05 were considered.
587
588 **Western blot analysis**
589 Cells were lysed in RIPA lysis buffer and the cell pellets were removed by centrifugation at
590 speed of 14,000 rpm for 30 min at 4 °C. Protein samples were prepared by mixing with
591 4× Laemmli Sample Buffer (Bio-Rad) supplemented with β-mercaptoethanol (Sigma-Aldrich)
592 and boiled at 95 °C for 10 min. After SDS-PAGE, the proteins were transferred from the gel to
593 polyvinylidene fluoride membranes (Bio-Rad). Membranes were blocked with 5% skim milk in
594 PBST (PBS supplemented with 0.1% Tween-20) for 1 h at room temperature and incubated
595 overnight with primary antibodies diluted in 5% skim milk in PBST at 4 °C. Membranes were
596 washed three times, 10 min each with PBST, incubated with secondary antibodies for one hour

at room temperature. Afterwards, membranes were washed three times, 10 min each with PBST. Positive immunostaining bands on the membranes were visualized using ECL Select Western Blotting Detection Reagent (Cytiva) and scanned by iBright 1500 imaging system (Invitrogen).

Growth kinetic analysis of virus

Confluent MDCK-SIAT1 or A549-Dual cells seeded in 24-well plates were washed with PBS once and infected with virus at the indicated MOI. After 1 h of adsorption, the viral inoculum was removed, and infected cells were washed twice with PBS, and then cultured in DMEM medium supplemented with either 0.5 μg/ml TPCK-treated trypsin for A549 or 1 μg/ml TPCK-treated trypsin for MDCK-SIAT1 at 37 °C. Supernatants were collected at the indicated time points by centrifugation at 13000 × g for 1 min to remove dead cells and stored at −80 °C until being titrated. Virus titers were determined by $TCID_{50}$ assay in MDCK-SIAT1 cells.

Quanti-luc luciferase assay

QUANTI-Luc Gold is a two-component reporter kit which contains: QUANTI-Luc Plus and QLC Stabilizer (Invivogen). A standard protocol according to manufacturer's instructions was followed. Quanti-Luc pouches were dissolved in sterile water together with QLC stabilizer. 20 μL of cell supernatant was added to a white opaque 96-well plate. 50 μL of QUANTI-Luc Gold assay solution was added to each well. The measurement was carried out immediately using BioTek synergy HTX multimode reader (Agilent).

Quanti-blue SEAP phosphatase assay

QUANTI-blue Gold is a two-component kit which contains: QB reagent and QB buffer (Invivogen). A standard protocol according to manufacturer's instructions was followed. The following protocol refers to the use of 96-well plates. 180 μL of QUANTI-Blue Solution was

623 dispensed per well into a flat-bottom 96-well plate. 20 µL of sample (supernatant of SEAP-
624 expressing cells) or negative control (cell culture medium) were added and incubated at 37 °C
625 for 6 h. Optical density (OD) at 620 nm was measured using a BioTek synergy HTX multimode
626 reader (Agilent).
627
628 **ISRE luciferase assay**
629 For luciferase assays, HEK293T cells were seeded in a 24-well plate at a density of 100,000
630 cells per well. The next day, cells were transiently transfected with pRL-TK and ISRE-Luc
631 reporter plasmids along with the indicated plasmids for 24 h. Cells were then treated overnight
632 with 2000 U/mL of universal Type I IFN (PBL Bioscience) and lysed at 24 h post treatment using
633 Passive Lysis Buffer (Promega). Samples were processed and luciferase activity was measured
634 using the Dual-Luciferase Assay System (Promega) according to the manufacturer's
635 instructions. Measurements were acquired by BioTek synergy HTX multimode reader (Agilent).
636 Firefly luciferase values were normalized to Renilla luciferase values.
637
638 **Caspase-3/7 activation assay**
639 Cells were stained with CellEventTM caspase-3/7 green flow cytometry assay kit (Thermo,
640 catalog #: C10427) according to the manufacturer's protocol. Data were acquired by a BD
641 Symphony A1 (BD Bioscience) flow cytometer, and the results were analyzed in FlowJo v10.8
642 software (BD Life Sciences).
643
644 **Influenza virus polymerase activity assay**
645 Dual luciferase activity reporter assays were performed to compare the polymerase activity of
646 RNP complexes in the presence of indicated NEP plasmid. RNP complex expression plasmids
647 composed of PB2, PB1, PA (125 ng each) and NP (250 ng), together with pYH-Luci reporter
648 plasmid (125 ng) and Renilla reporter plasmid (10 ng) were co-transfected into HEK293T cells.

649 At 24 h after transfection, cells were lysed, and the luciferase activity was measured using the
650 Dual-Luciferase reporter assay system (Promega). Measurements were acquired by a BioTek
651 synergy HTX multimode reader (Agilent).
652
653 **Quantification and statistical analysis**
654 All the statistical analyses have been performed using Prism 9 Graph Pad Software. Two-tailed
655 Student's unpaired t-test was performed to compare between two populations of data (e.g.,
656 control and sample) whereas two-way ANOVA was applied for multi sample comparisons. All
657 data generated were from independent biological replicates where n ≥ 3, each measured in
658 technical duplicates or triplicates. Results have been presented as means ± standard deviation
659 (SD) or standard error of the mean (SEM).
660
661 **Data availability**
662 Raw sequencing data have been submitted to the NIH Short Read Archive under accession
663 numbers: BioProject PRJNA1083715 and BioProject PRJNA1110270.
664
665 **Code availability**
666 Custom python scripts for all analyses have been deposited to:
667 https://github.com/nicwulab/WSN_NEP_DMS.
668
669 **REFERENCES**

670 1. Wu NC, Wilson IA. Structural insights into the design of novel anti-influenza therapies. *Nat*
671 *Struct Mol Biol* **25**, 115-121 (2018).
672
673 2. Chua MA, Schmid S, Perez JT, Langlois RA, Tenoever BR. Influenza A virus utilizes
674 suboptimal splicing to coordinate the timing of infection. *Cell Rep* **3**, 23-29 (2013).
675

3. Haye K, Burmakina S, Moran T, Garcia-Sastre A, Fernandez-Sesma A. The NS1 protein of a human influenza virus inhibits type I interferon production and the induction of antiviral responses in primary human dendritic and respiratory epithelial cells. *J Virol* **83**, 6849-6862 (2009).

4. O'Neill RE, Talon J, Palese P. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO J* **17**, 288-296 (1998).

5. Garcia-Sastre A*, et al.* Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. *Virology* **252**, 324-330 (1998).

6. Neumann G, Hughes MT, Kawaoka Y. Influenza A virus NS2 protein mediates vRNP nuclear export through NES-independent interaction with hCRM1. *EMBO J* **19**, 6751-6758 (2000).

7. O'Neill RE, Talon J, Palese P. The influenza virus NEP (NS2 protein) mediates the nuclear export of viral ribonucleoproteins. *EMBO J* **17**, 288-296 (1998).

8. Huang S*, et al.* A second CRM1-dependent nuclear export signal in the influenza A virus NS2 protein contributes to the nuclear export of viral ribonucleoproteins. *J Virol* **87**, 767-778 (2013).

9. Akarsu H*, et al.* Crystal structure of the M1 protein-binding domain of the influenza A virus nuclear export protein (NEP/NS2). *EMBO J* **22**, 4646-4655 (2003).

10. Paterson D, Fodor E. Emerging roles for the influenza A virus nuclear export protein (NEP). *PLoS Pathog* **8**, e1003019 (2012).

11. Zhu Z, Fodor E, Keown JR. A structural understanding of influenza virus genome replication. *Trends Microbiol* **31**, 308-319 (2023).

12. Poon LL, Pritlove DC, Fodor E, Brownlee GG. Direct evidence that the poly (A) tail of influenza A virus mRNA is synthesized by reiterative copying of a U track in the virion RNA template. *J Virol* **73**, 3473-3476 (1999).

13. Te Velthuis AJ, Grimes JM, Fodor E. Structural insights into RNA polymerases of negative-sense RNA viruses. *Nat Rev Microbiol* **19**, 303-318 (2021).

14. Phan T, Fay EJ, Lee Z, Aron S, Hu W-S, Langlois RA. Segment-specific kinetics of mRNA, cRNA, and vRNA accumulation during influenza virus infection. *J Virol* **95**, e02102-02120 (2021).

15. Hatada E, Hasegawa M, Mukaigawa J, Shimizu K, Fukuda R. Control of influenza virus gene expression: quantitative analysis of each viral RNA species in infected cells. *J Biochem* **105**, 537-546 (1989).

16. Bullido R, Gomez-Puertas P, Saiz MJ, Portela A. Influenza A virus NEP (NS2 protein) downregulates RNA synthesis of model template RNAs. *J Virol* **75**, 4912-4917 (2001).

17. Chua MA, Schmid S, Perez JT, Langlois RA, tenOever BR. Influenza A virus utilizes suboptimal splicing to coordinate the timing of infection. *Cell Rep* **3**, 23-29 (2013).

18. Robb NC, Smith M, Vreede FT, Fodor E. NS2/NEP protein regulates transcription and replication of the influenza virus RNA genome. *J Gen Virol* **90**, 1398-1407 (2009).

19. Perez JT*, et al.* Influenza A virus-generated small RNAs regulate the switch from transcription to replication. *PNAS* **107**, 11525-11530 (2010).

20. Nilsson BE, Te Velthuis AJW, Fodor E. Role of the PB2 627 domain in influenza A virus polymerase function. *J Virol* **91**, e02467-02416 (2017).

21. Sun L*, et al.* The SUMO-interacting motif in NS2 promotes adaptation of avian influenza virus to mammals. *Sci Adv* **9**, eadg5175 (2023).

22. Manz B, Brunotte L, Reuther P, Schwemmle M. Adaptive mutations in NEP compensate for defective H5N1 RNA replication in cultured human cells. *Nat Commun* **3**, 802 (2012).

23. Reuther P*, et al.* Adaptive mutations in the nuclear export protein of human-derived H5N1 strains facilitate a polymerase activity-enhancing conformation. *J Virol* **88**, 263-271 (2014).

24. Lommer BS, Luo M. Structural plasticity in influenza virus protein NS2 (NEP). *J Biol Chem* **277**, 7108-7117 (2002).

25. Basler CF*, et al.* Sequence of the 1918 pandemic influenza virus nonstructural gene (NS) segment and characterization of recombinant viruses bearing the 1918 NS genes. *PNAS* **98**, 2746-2751 (2001).

26. Neumann G*, et al.* Generation of influenza A viruses entirely from cloned cDNAs. *PNAS* **96**, 9345-9350 (1999).

27. Song W*, et al.* The K526R substitution in viral protein PB2 enhances the effects of E627K on influenza virus replication. *Nature Commun* **5**, 5509 (2014).

28. Brunotte L, Flies J, Bolte H, Reuther P, Vreede F, Schwemmle M. The nuclear export protein of H5N1 influenza A viruses recruits Matrix 1 (M1) protein to the viral ribonucleoprotein to mediate nuclear export. *J Biol Chem* **289**, 20067-20077 (2014).

29. Hirohama M, Yamashita S, Asaka MN, Kuroki T, Kawaguchi A. Intramolecular interaction of NEP regulated by CRM1 ensures the unidirectional transport of M1 for the nuclear export of influenza viral ribonucleoprotein. *Front Virol* **3**, 1232906 (2023).

30. Fujii Y, Goto H, Watanabe T, Yoshida T, Kawaoka Y. Selective incorporation of influenza virus RNA segments into virions. *PNAS* **100**, 2002-2007 (2003).

31. Iwatsuki-Horimoto K, Horimoto T, Fujii Y, Kawaoka Y. Generation of influenza A virus NS2 (NEP) mutants with an altered nuclear export signal sequence. *J Virol* **78**, 10149-10155 (2004).

32. Odagiri T, Tominaga K, Tobita K, Ohta S. An amino acid change in the non-structural NS2 protein of an influenza A virus mutant is responsible for the generation of defective interfering (DI) particles by amplifying DI RNAs and suppressing complementary RNA synthesis. *J Gen Virol* **75**, 43-53 (1994).

33. Zhang L, Wang Y, Shao Y, Guo J, Gao GF, Deng T. Fine regulation of influenza virus RNA transcription and replication by stoichiometric changes in viral NS1 and NS2 proteins. *J Virol* **97**, e0033723 (2023).

34. Te Velthuis AJW, Long JS, Barclay WS. Assays to measure the activity of influenza virus polymerase. *Methods Mol Biol* **1836**, 343-374 (2018).

35. Martin BE, Harris JD, Sun J, Koelle K, Brooke CB. Cellular co-infection can modulate the efficiency of influenza A virus production and shape the interferon response. *PLoS Pathogens* **16**, e1008974 (2020).

36. Zhang B*, et al.* H1N1 influenza A virus protein NS2 inhibits innate immune response by targeting IRF7. *Viruses* **14**, 2411 (2022).

37. Lebeau G*, et al.* The efficient antiviral response of A549 cells is enhanced when mitochondrial respiration is promoted. *Pathogens* **11**, 1168 (2022).

38. Shim JM, Kim J, Tenson T, Min J-Y, Kainov DE. Influenza virus infection, interferon response, viral counter-response, and apoptosis. *Viruses* **9**, 223 (2017).

39. Lee S, Hirohama M, Noguchi M, Nagata K, Kawaguchi A. Influenza A virus infection triggers pyroptosis and apoptosis of respiratory epithelial cells through the type I interferon signaling pathway in a mutually exclusive manner. *J Virol* **92**, 00396-00318 (2018).

40. El-Sayed I, Bassiouny K, Nokaly A, Abdelghani AS, Roshdy W. Influenza A virus and influenza B virus can induce apoptosis via intrinsic or extrinsic pathways and also via NF-κB in a time and dose dependent manner. *Bioche Res Int* **2016**, 1738237 (2016).

41. Laghlali G, Lawlor KE, Tate MD. Die another way: interplay between influenza A virus, inflammation and cell death. *Viruses* **12**, 401 (2020).

42. Herold S, Ludwig S, Pleschka S, Wolff T. Apoptosis signaling in influenza virus propagation, innate host defense, and lung injury. *J Leukoc Biol* **92**, 75-82 (2012).

43. Zhang L*, et al.* Twenty natural amino acid substitution screening at the last residue 121 of influenza A virus NS2 protein reveals the critical role of NS2 in promoting virus genome replication by coordinating with viral polymerase. *J Virol* **98**, e01166-01123 (2024).

44. Liu Y-C*, et al.* NS2 is a key determinant of compatibility in reassortant avian influenza virus with heterologous H7N9-derived NS segment. *Virus Res* **324**, 199028 (2023).

45. Huang X*, et al.* An NS-segment exonic splicing enhancer regulates influenza A virus replication in mammalian cells. *Nat Commun* **8**, 14751 (2017).

46. Richardson J, Akkina R. NS2 protein of influenza virus is found in purified virus and phosphorylated in infected cells. *Arch Virol* **116**, 69-80 (1991).

47. Yasuda J, Nakada S, Kato A, Toyoda T, Ishihama A. Molecular assembly of influenza virus: association of the NS2 protein with virion matrix. *Virology* **196**, 249-255 (1993).

48. Fernandes JD*, et al.* Functional segregation of overlapping genes in HIV. *Cell* **167**, 1762-1773. e1712 (2016).

49. Qian W*, et al.* The C-terminal effector domain of non-structural protein 1 of influenza A virus blocks IFN-β production by targeting TNF receptor-associated factor 3. *Front Immunol* **8**, 779 (2017).

50. Turianová L, Lachová V, Beňová K, Kostrábová A, Betáková T. Influenza A virus lacking the effector and C terminal domains of NS1 protein induces cytokines associated with high pathogenicity in mice. *Acta Virol* **64**, 78-87 (2020).

51. Wei K*, et al.* Influenza A virus acquires enhanced pathogenicity and transmissibility after serial passages in swine. *J Virol* **88**, 11981-11994 (2014).

52. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. *PNAS* **108**, 9530-9535 (2011).

53. Wu NC*, et al.* High-throughput profiling of influenza A virus hemagglutinin gene at single-nucleotide resolution. *Sci Rep* **4**, 4942 (2014).

54. Zhang TH, Wu NC, Sun R. A benchmark study on error-correction by read-pairing and tag-clustering in amplicon-based deep sequencing. *BMC Genomics* **17**, 108 (2016).

55. Koster J, Rahmann S. Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* **34**, 3600 (2018).

56. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**, 10-12 (2011).

57. Magoč T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957-2963 (2011).

58. Cock PJ*, et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422 (2009).

59. Dobin A*, et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

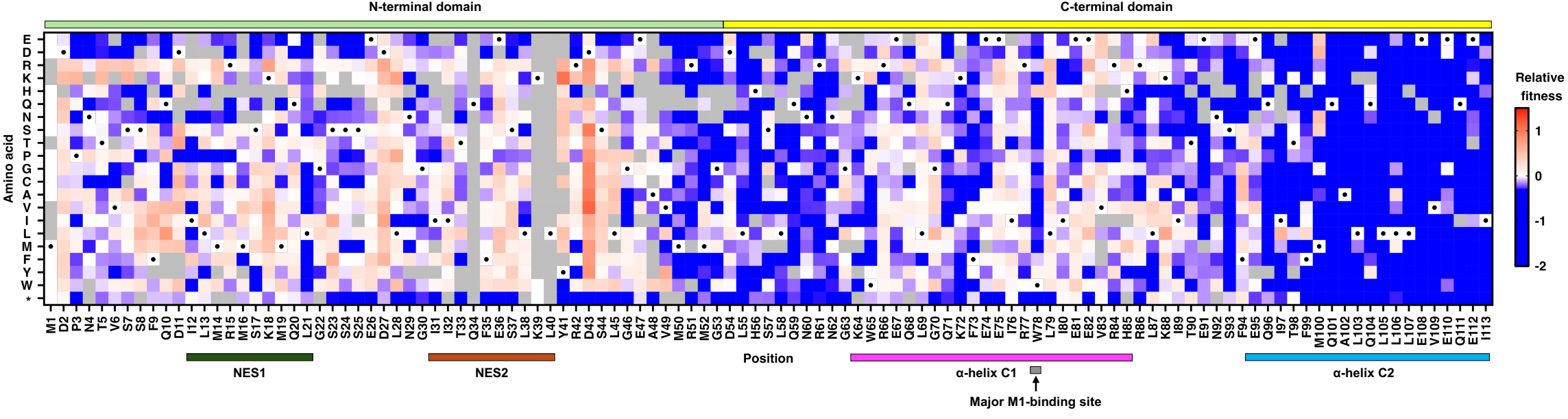
60. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).

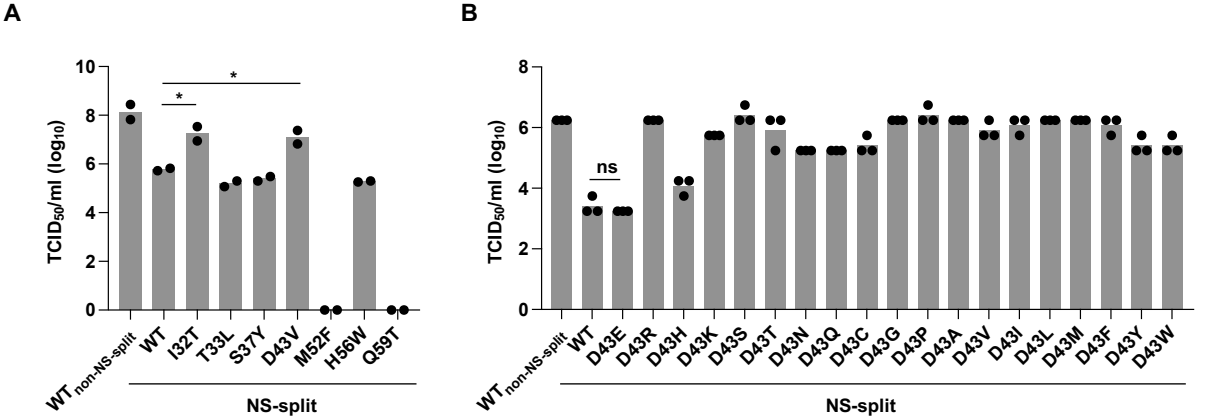
61. Hoffmann E, Stech J, Guan Y, Webster R, Perez D. Universal primer set for the full-length amplification of all influenza A viruses. *Arch Virol* **146**, 2275-2289 (2001).

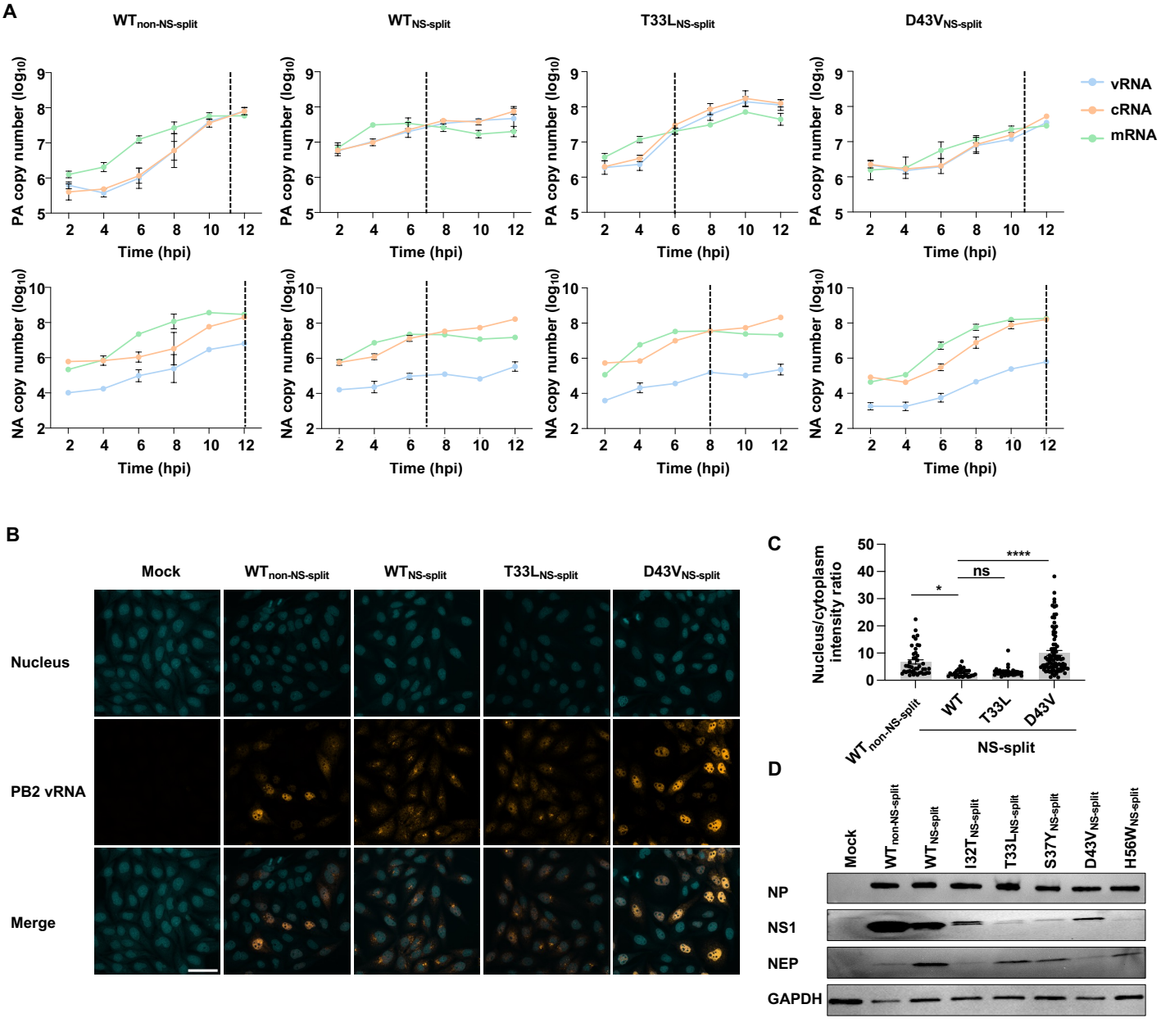
62. Desselberger U, Racaniello VR, Zazra JJ, Palese P. The 3'and 5'-terminal sequences of influenza A, B and C virus RNA segments are highly conserved and show partial inverted complementarity. *Gene* **8**, 315-328 (1980).

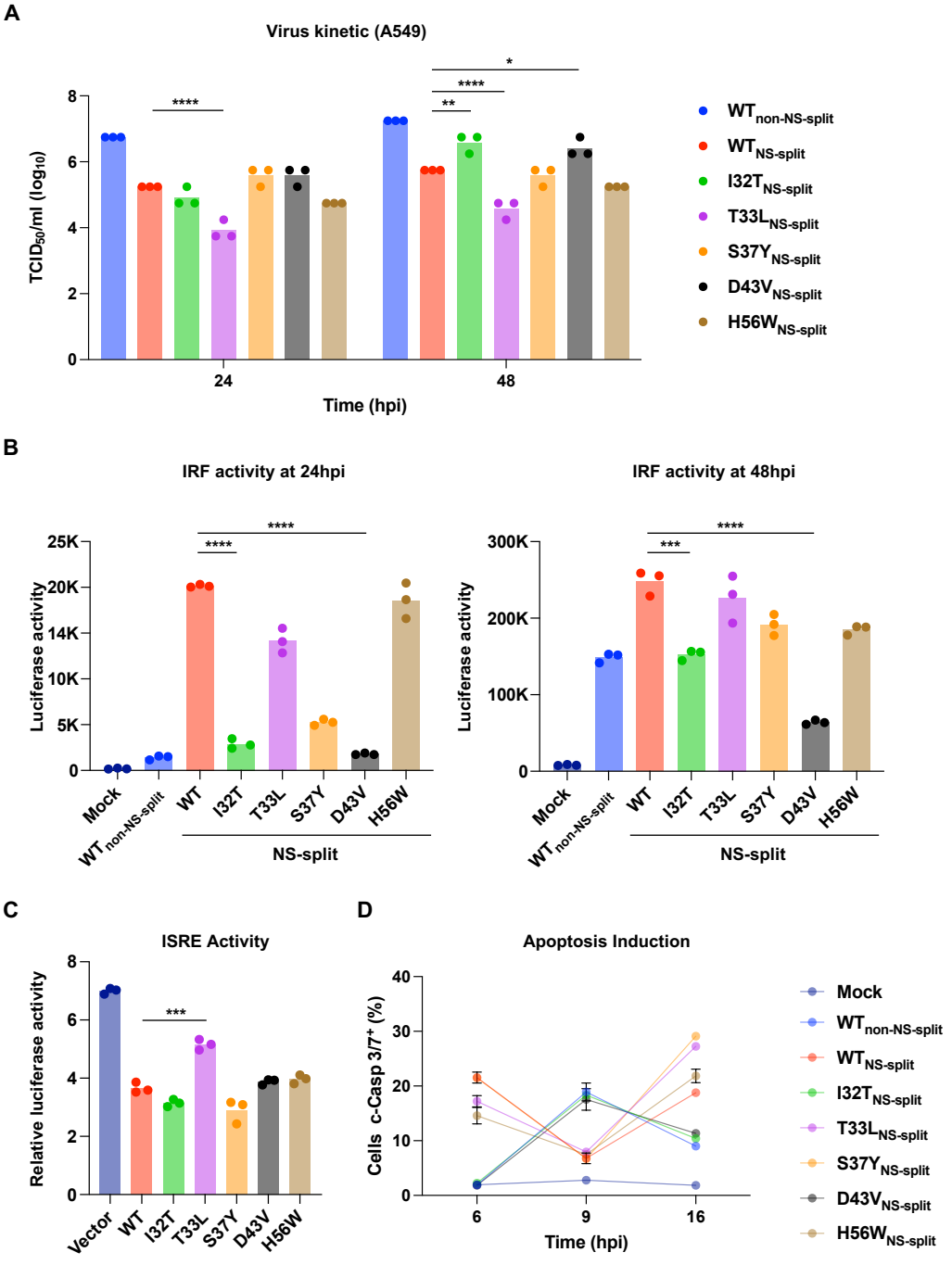
63. Schwarzkopf M*, et al.* Hybridization chain reaction enables a unified approach to multiplexed, quantitative, high-resolution immunohistochemistry and in situ hybridization. *Development* **148**, (2021).

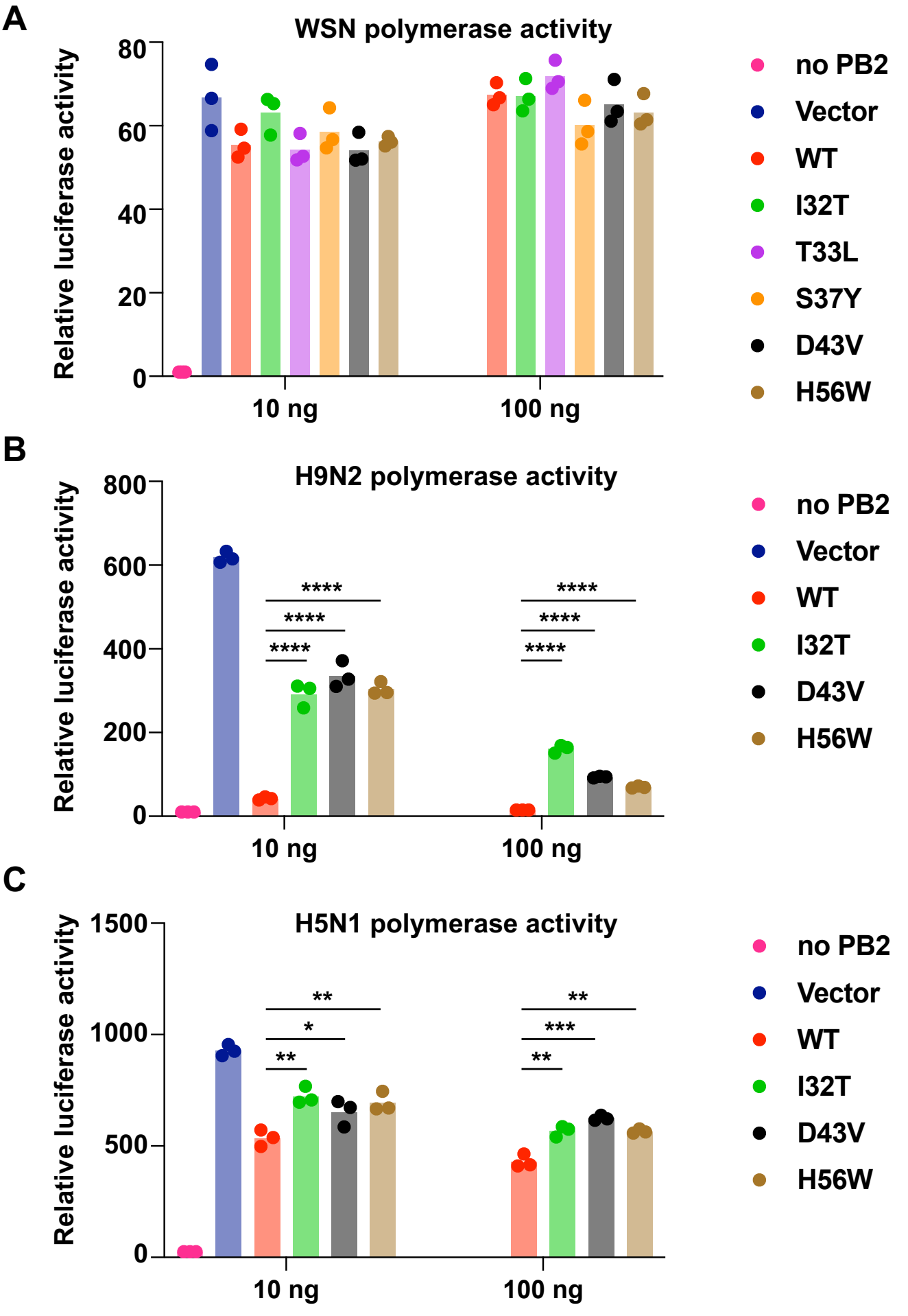
64. Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A. CellProfiler 4: improvements in speed, utility and usability. *BMC bioinformatics* **22**, 1-11 (2021).


N-terminal domain
C-terminal domain
Amino acid
E
D
R
K
H
Q
N
S
T
P
G
C
A
V
I
L
M
F
Y
W
*
M1 D2 P3 N4 T5 V6 S7 S8 F9 Q10 D11 I12 L13 M14 R15 M16 S17 K18 M19 Q20 L21 G22 S23 S24 S25 E26 D27 L28 N29 G30 I31 I32 T33 Q34 F35 E36 S37 L38 K39 L40 Y41 R42 D43 S44 L45 G46 E47 A48 V49 M50 R51 M52 G53 D54 L55 H56 S57 L58 Q59 N60 R61 N62 G63 K64 W65 R66 E67 Q68 L69 G70 Q71 K72 F73 E74 E75 I76 R77 W78 L79 I80 E81 E82 V83 R84 H85 R86 L87 K88 I89 T90 E91 N92 S93 F94 E95 Q96 I97 T98 F99 M100 Q101 A102 L103 Q104 L105 L106 L107 E108 V109 E110 Q111 E112 I113
NES1
NES2
Position
α-helix C1
α-helix C2
Major M1-binding site
Relative fitness
1
0
-1
-2


A
B
TCID50/ml (log10)
10
8
6
4
2
0
*
*
ns
WT non-NS-split
WT
I32T
T33L
S37Y
D43V
M52F
H56W
Q59T
D43E
D43R
D43H
D43K
D43S
D43T
D43N
D43Q
D43C
D43G
D43P
D43A
D43V
D43I
D43L
D43M
D43F
D43Y
D43W
NS-split


A
WTnon-NS-split
WTNS-split
T33LNS-split
D43VNS-split
vRNA
cRNA
mRNA
PA copy number (log10)
NA copy number (log10)
Time (hpi)
B
Mock
WTnon-NS-split
WTNS-split
T33LNS-split
D43VNS-split
Nucleus
PB2 vRNA
Merge
C
Nucleus/cytoplasm intensity ratio
*
ns
****
WTnon-NS-split
WT
T33L
D43V
NS-split
D
Mock
WTnon-NS-split
WTNS-split
I32TNS-split
T33LNS-split
S37YNS-split
D43VNS-split
H56WNS-split
NP
NS1
NEP
GAPDH


A
Virus kinetic (A549)
TCID50/ml (log10)
Time (hpi)
24
48
WTnon-NS-split
WTNS-split
I32TNS-split
T33LNS-split
S37YNS-split
D43VNS-split
H56WNS-split
B
IRF activity at 24hpi
IRF activity at 48hpi
Luciferase activity
Mock
WTnon-NS-split
WT
I32T
T33L
S37Y
D43V
H56W
NS-split
C
ISRE Activity
Relative luciferase activity
Vector
D
Apoptosis Induction
Cells c-Casp 3/7+ (%)


A
WSN polymerase activity
Relative luciferase activity
80
60
40
20
0
10 ng
100 ng
no PB2
Vector
WT
I32T
T33L
S37Y
D43V
H56W
B
H9N2 polymerase activity
Relative luciferase activity
800
600
400
200
0
****
****
****
****
****
****
10 ng
100 ng
no PB2
Vector
WT
I32T
D43V
H56W
C
H5N1 polymerase activity
Relative luciferase activity
1500
1000
500
0
**
*
**
**
***
**
10 ng
100 ng
no PB2
Vector
WT
I32T
D43V
H56W