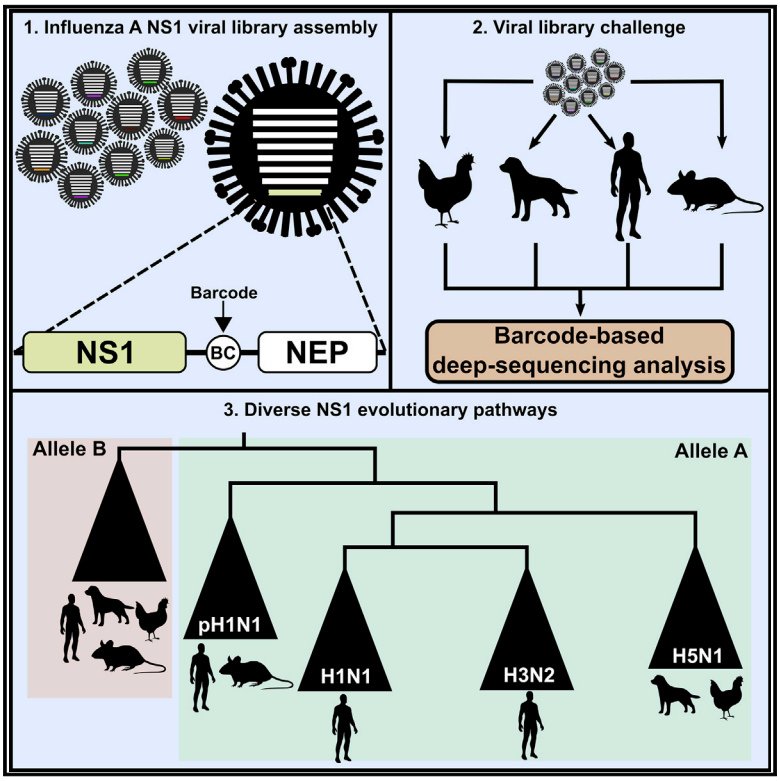
Article

Cell Reports

Viral Fitness Landscapes in Diverse Host Species Reveal Multiple Evolutionary Lines for the NS1 Gene of Influenza A Viruses

Graphical Abstract



Highlights

- Influenza A NS1-mediated host tropism is versatile and constantly evolving
- Phylogenetically related NS1 can display divergent phenotypic profiles
- Avian-origin allele B NS1 efficiently replicates in a range of hosts
- The library allows the study of NS1 fitness contribution within a viral population

Authors

Raquel Muñoz-Moreno, Carles Martínez-Romero, Daniel Blanco-Melo, ..., Florian Krammer, Benjamin R. tenOever, Adolfo García-Sastre

Correspondence

adolfo.garcia-sastre@mssm.edu

In Brief

Muñoz-Moreno et al. report that influenza A virus NS1 undergoes diverse and unpredictable evolutionary pathways based on its different phylogenetic lineages. A high-throughput approach using a barcoded library is used to test the interactions between NS1-recombinant viruses and to study their preference for specific or multiple hosts.

Muñoz-Moreno et al., 2019, Cell Reports *29*, 3997–4009
December 17, 2019
https://doi.org/10.1016/j.celrep.2019.11.070

CellPress

Cell Reports
Article

Viral Fitness Landscapes in Diverse Host Species Reveal Multiple Evolutionary Lines for the NS1 Gene of Influenza A Viruses

Raquel Muñoz-Moreno,[1,2,8] Carles Martínez-Romero,[1,2,8] Daniel Blanco-Melo,[1] Christian V. Forst,[3] Raffael Nachbagauer,[1] Asiel Arturo Benitez,[1,2,9] Ignacio Mena,[1,2] Sadaf Aslam,[1] Vinod Balasubramaniam,[1,2,4] Ilseob Lee,[1,2,10] Maryline Panis,[1,2] Juan Ayllón,[1,2] David Sachs,[3] Man-Seong Park,[5] Florian Krammer,[1] Benjamin R. tenOever,[1,2] and Adolfo García-Sastre[1,2,6,7,11,*]

[1]Department of Microbiology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[2]Global Health and Emerging Pathogens Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[3]Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[4]Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, 47500 Bandar Sunway, Malaysia
[5]Department of Microbiology, Institute for Viral Diseases, College of Medicine, Korea University, Seoul 02841, Republic of Korea
[6]Department of Medicine, Division of Infectious Diseases, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[7]The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[8]These authors contributed equally
[9]Present address: Regeneron Pharmaceuticals, 777 Old Saw Mill River Road, Tarrytown, NY 10591, USA
[10]Present address: Veterinary Research Division, Animal and Plant Quarantine Agency, Gimcheon-si, Gyeongsangbuk-do 39660. Republic of Korea
[11]Lead Contact
*Correspondence: adolfo.garcia-sastre@mssm.edu
https://doi.org/10.1016/j.celrep.2019.11.070

SUMMARY

Influenza A viruses (IAVs) have a remarkable tropism in their ability to circulate in both mammalian and avian species. The IAV NS1 protein is a multifunctional virulence factor that inhibits the type I interferon host response through a myriad of mechanisms. How NS1 has evolved to enable this remarkable property across species and its specific impact in the overall replication, pathogenicity, and host preference remain unknown. Here we analyze the NS1 evolutionary landscape and host tropism using a barcoded library of recombinant IAVs. Results show a surprisingly great variety of NS1 phenotypes according to their ability to replicate in different hosts. The IAV NS1 genes appear to have taken diverse and random evolutionary pathways within their multiple phylogenetic lineages. In summary, the high evolutionary plasticity of this viral protein underscores the ability of IAVs to adapt to multiple hosts and aids in our understanding of its global prevalence.

INTRODUCTION

Several influenza A virus (IAV) subtypes and strains co-circulate worldwide in diverse hosts, a subset of which causes seasonal epidemics every year in humans. Constant evolution of these heterogeneous viral populations is driven by small, single-point mutation changes or by the result of reassortment within the same host due to co-infection events (Treanor, 2004). These evolutionary events can lead to the emergence of new strains with pandemic potential in humans. Thus, developing ways to efficiently and accurately detect viral genotypic changes and predict fitness-based phenotypic outcomes within a viral population is imperative for predicting future pandemics. Influenza viruses are part of the negative single-stranded RNA family *Orthomyxoviridae*, with an approximately 13-kb-long segmented genome (Palese and Shaw, 2007). Segment 8 (NS) is the smallest and encodes for the nonstructural protein NS1, as well as for the NEP protein generated through an alternative splicing event. NS1 is a significant contributor to the pathogenic potential of IAVs and has become an important research topic in influenza research, because it displays multiple mechanisms to antagonize the interferon (IFN)-mediated host antiviral response, some of which are viral strain and host species specific (García-Sastre et al., 1998; Greenspan et al., 1988; Kochs et al., 2007). NS1 is expressed in the nucleus and in the cytoplasm of infected cells, interacting functionally with different cellular factors. For instance, the N-terminal RNA binding domain of NS1 binds to double-stranded RNA (dsRNA) molecules, thus blocking the activation of the antiviral enzyme OAS (Min and Krug, 2006). NS1 also binds and sequesters dsRNA from activating RIG-I Like Receptors (RLRs), Protein Kinase R (PKR), and MDA5, which are triggers of the host innate immune response (Benitez et al., 2015; Hatada and Fukuda, 1992; Liu et al., 1997). In addition, RIG-I and PKR are specifically inhibited by NS1 through direct binding to these proteins or to proteins involved in their activation (Bergmann et al., 2000; Guo et al., 2007; Li et al., 2006). NS1 also blocks mRNA maturation and export from the nucleus to the cytoplasm, thus inhibiting cellular gene expression (Nemeroff et al., 1998; Satterly et al., 2007). Other host factors interacting with NS1 have been described, resulting in the regulation of multiple host pathways related to innate immunity, cell death, transcription, and mRNA processing

(Ayllon and García-Sastre, 2015). Interestingly, due to potential redundancy among the different NS1-mediated effects, several influenza strains contain NS1 phenotypes that do not possess all functions described to date (Hale et al., 2010). Furthermore, some host proteins that interact with NS1 do not share identical sequences across species. This opens the possibility for differences in virus-host protein interactions according to the viral strain origin of the NS1 and the host species origin of the cellular factor. Such differences are likely to favor or to restrict the preferred host for viral replication.

Based on amino acid homology, NS segments can be classified into two distinct gene pools: allele A and allele B (Ludwig et al., 1991). The allele A group includes NS variants from both avian and mammal origins, whereas the allele B group mainly contains avian-origin strains that have been recently reported to be also capable of supporting infection in mammalian hosts (Turnbull et al., 2016). Sequence identity within each group is typically above 90%, but it is only around 60% between both alleles (Munir et al., 2011; Zohari et al., 2010). In this study, we have selected 48 allele A and 9 allele B NS1 phenotypes representing the sequence space of the NS1 present in IAV sequence databases and compared the ability of each NS1 to contribute to the replication of IAVs in multiple host substrates.

RESULTS

Current findings of NS1 and its role in viral replication have been studied in the context of single viral infections. Given that influenza virus strains are continuously co-circulating and evolving through selective pressure, studying the spatio-temporal evolution of NS1 in a wider context may provide insights in its contributions to IAV replication and tropism. To address this, we used the influenza H1N1 (A/Puerto Rico/8/1934) virus background to generate a library of recombinant IAVs expressing a selection of NS1 proteins within a modified NS segment. This strategy allows us to determine the abundance of each virus within the library through a barcode-based next-generation sequencing (NGS) technology (Varble et al., 2014). The splicing mechanism of the NS segment was altered to separate both NS1 and NEP open reading frames (ORFs), allowing changes in the NS1 sequence without affecting the NEP sequence. In addition, we inserted a genetically neutral 22-nucleotide barcode in the non-coding region between the NS1 and the NEP ORFs (Figure S1A). This strategy enables us to track individual viral fitness within the library based on the relative number of barcode reads after replication in different hosts.

After confirming that expression of the modified NS segment did not significantly affect viral replication (Figure S1B), we validated our approach to study NS1-mediated fitness in multiple hosts using a mixture of 4 recombinant IAVs containing the following NS1 sequences: A/Shanghai/02/2013 from an avian H7N9 strain, A/Udorn/1972 from a human H3N2 strain, A/Puerto Rico/8/1934 (PR8) from a human mouse-adapted H1N1 strain, and the loss-of-function mutant PR8-R38A/K41A, which is unable to bind to dsRNA (Donelan et al., 2003; Talon et al., 2000). Several models of influenza virus infection—Madin-Darby canine kidney (MDCK) cells, embryonated chicken eggs, and 8-week-old C57BL/6 mice—were infected with a mix containing either 50 or 100 plaque-forming units (PFUs) per virus. Total barcode reads per virus were normalized to the initial proportion present on the viral mix before infection (input) after 48 h of replication and plotted in viral progression diagrams (Figure S1C). As anticipated, barcode reads of PR8-R38A/K41A NS1 severely decreased compared with input. This was also the case for the NS1 of A/Udorn/1972, although to a lesser extent. However, levels of wild-type (WT) PR8 notably increased, followed by A/Shanghai/02/2013. These results support our barcode-based strategy as a tool to study individual viral fitness within a library of IAVs only differing in their NS1 sequences. Because generally similar profiles were obtained between replicates at both concentrations tested in all three infection models, a dose of 100 PFU/virus was selected for further experiments with our full library of recombinant IAVs.

We next selected various NS1 variants from fully sequenced IAV strains based on their relative phylogenetic distance, their host, and their year of isolation. To maximize functional representation, we also took into consideration the inclusion of NS1 sequences with different structural and functional sequence feature variant types (SFVTs) (Noronha et al., 2012). Table 1 contains our final selection of 56 NS1 sequences from different strains, hosts, countries, and years. Figure 1 shows the position of our selected NS1 sequences within a global NS1 phylogenetic tree. Their widespread distribution among different clades confirms that our curated viral library is representative of the NS1 sequence space in circulating IAVs from 1954 to 2013. The selected NS1 sequences were assembled into split NS segments and assigned to unique barcode sequences. Recombinant viruses were rescued by reverse genetics in an influenza A/Puerto Rico/8/1934 virus background (Fodor et al., 1999; Neumann et al., 1999). While each virus contained a unique NS1 sequence, the rest of the viral genome remained constant within the library.

Importantly, each NS1 sequence in our final virus library was represented by two recombinant viruses, each with a unique barcode. This was implemented as an internal control of variability, because viruses that carry the same NS1 but have different barcodes should behave similarly within the library. Moreover, this duplication approach is used to account for possible differences in amplification due to the different nucleotide composition among the barcode sequences. Recombinant viruses were individually rescued, grown in MDCK cells, titrated, and fully sequenced to confirm that no additional changes were generated during the rescue and that our virus stocks were generally free of defective interfering RNAs. Finally, all individual viruses were mixed in a common library preparation using identical numbers of infectious units. Then, MDCK cells, embryonated chicken eggs, and mice were infected with the library, and 48 h later, barcode abundance for each virus was analyzed and normalized to the original library (input) (Figure 2). Viruses containing phylogenetically related NS1 behaved similarly in all three hosts tested, and no major differences were found in barcode abundance among duplicated NS1 viruses. However, overall fitness diversity was observed due to the NS1 species-specific tropism within the library (Figure 2A). For example, avian H5N1 NS1-containing viruses replicated poorly in mouse lungs compared with MDCK supernatants and egg allantoic fluid. In

Table 1. List of NS1 Strains Included in the Viral Library

Strain Name	Subtype	Species	Country	Accession No.	Allele
A/Peru/PER271/2011	H3N2	human	Peru	CY162412	A
A/Singapore/H2011.447/2011	H3N2	human	Singapore	KF014773	A
A/Canterbury/204/2005	H3N2	human	New Zealand	CY008376	A
A/Scotland/72/2003	H3N2	human	United Kingdom	CY088106	A
A/Finland/170/2003	H3N2	human	Finland	CY114353	A
A/Canterbury/01/2002	H3N2	human	New Zealand	CY007591	A
A/South Australia/62/2000	H3N2	human	Australia	CY016719	A
A/Auckland/599/2000	H3N2	human	New Zealand	CY023038	A
A/Finland/381/1995	H3N2	human	Finland	CY114185	A
A/Umea/2000/1992	H3N2	human	Sweden	CY113945	A
A/Geneva/5366/1991	H3N2	human	Switzerland	CY113569	A
A/Stockholm/10/1985	H3N2	human	Sweden	CY113345	A
A/Rotterdam/577/1980	H3N2	human	Netherlands	CY112357	A
A/Bilthoven/334/1975	H3N2	human	Netherlands	CY113153	A
A/Hong Kong/49/1974	H3N2	human	Hong Kong	CY006911	A
A/Udorn/1972	H3N2	human	Russia	CY009640	A
A/Beijing/1/1968	H3N2	human	China	CY008160	A
A/Cottbus/1/1964	H2N2	human	Germany	CY032265	A
A/Boston/6/2009	H1N1	human	USA	CY089039	A
A/Memphis/19/1983	H1N1	human	USA	CY012444	A
A/Albany/20/1978	H1N1	human	USA	CY021801	A
A/Malaysia/54	H1N1	human	Malaysia	CY009344	A
A/Puerto Rico/8/1934	H1N1	human	Puerto Rico	CY148247	A
A/Anhui/2/2005	H5N1	human	China	CY098581	A
A/Duck/Guangdong/23/2004	H5N1	bird	China	HM172300	A
A/Thailand/676/2005	H5N1	human	Thailand	DQ360842	A
A/Vietnam/1203/2004	H5N1	human	Vietnam	EF541456	A
A/Chicken/Moscow/2/2007	H5N1	bird	Russia	EF474446	A
A/Cygnus olor/Croatia/1/2005	H5N1	bird	Croatia	CY016823	A
A/Duck/Fujian/7844/2007	H6N6	bird	China	CY110576	A
A/Broiler duck/Korea/H32/2014	H5N8	bird	South Korea	KJ508920	A
A/Mallard/Republic of Georgia/12/2011	H10N4	bird	Georgia	CY185509	A
A/Blue-winged teal/Guatemala/CIP049-14/2010	H8N4	bird	Guatemala	CY096692	A
A/Shorebird/Delaware Bay/107/2009	H10N7	bird	USA	CY137774	A
A/Shanghai/02/2013	H7N9	human	China	KF021601	A
A/Chicken/Rizhao/651/2013	H9N2	bird	China	KF260213	A
A/Hong Kong/97/98	H5N1	human	Hong Kong	AF256188	A
A/Camel/Mongolia/335/2012	H3N8	camel	Mongolia	CY164124	A
A/Equine/Xinjiang/2/2007	H3N8	horse	China	EU794555	A
A/Canine/Colorado/30604/2006	H3N8	dog	USA	CY067386	A
A/Equine/Berlin/1/1989	H3N8	horse	Germany	CY032417	A
A/Equine/Uruguay/1063/1976	H7N7	horse	Uruguay	CY036891	A
A/Athens/INS419/2010	H1N1	human	Greece	CY071307	A
A/Elephant seal/California/1/2010	H1N1	sea mammal	USA	JX865426	A
A/Alaska/01/2010	H1N1	human	USA	KC781390	A
A/Akita/1/2009	H1N1	human	Japan	GQ365414	A
A/California/07/2009	H1N1	human	USA	FJ969538	A

(*Continued on next page*)

Table 1. *Continued*

Strain Name	Subtype	Species	Country	Accession No.	Allele
A/Swine/Manitoba/01643/2007	H3N2	pig	Canada	CY158669	A
A/Mallard/Netherlands/1/2010	H2N3	bird	Netherlands	CY122312	B
A/Duck/Tasmania/277/2007	H7N2	bird	Australia	CY033165	B
A/Mallard/Sweden/79367/2008	H4N6	bird	Sweden	CY165389	B
A/Duck/Yangzhou/02/2005	H8N4	bird	China	EF061119	B
A/American black duck/New Brunswick/00477/2010	H10N6	bird	Canada	CY138937	B
A/Mallard/California/1154/2010	H4N6	bird	USA	CY125954	B
A/Mallard/Nova Scotia/00088/2010	H1N1	bird	Canada	CY138564	B
A/Green-winged teal/Alberta/228/1985	H7N3	bird	Canada	CY185733	B
A/Duck/New York/21211-6/2005	H7N2	bird	USA	CY022849	B

Information regarding the subtype, host, country of origin, NCBI accession number, and allele group is specified for each protein. NS1 strains are listed based on their phylogenetic relationship (see Figures 2, 5, S2, and S6).

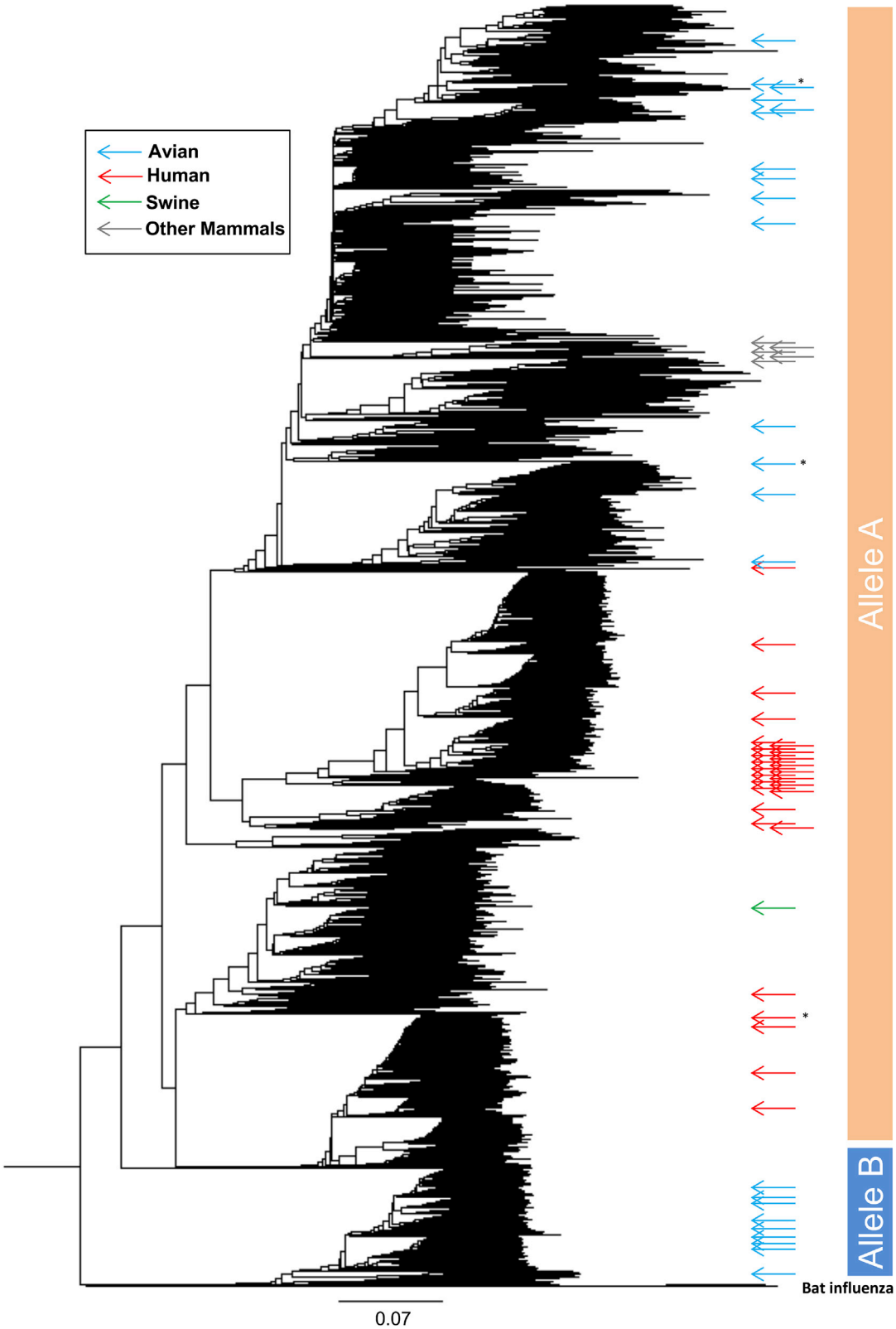
contrast, viruses carrying human H3N2 NS1 were underrepresented in all three models tested, suggesting poor adaptation to canine, chicken, and murine hosts. Conversely, viruses containing mammalian non-human NS1 (H3N8) were overrepresented in MDCK supernatants and in chicken allantoic fluid. Interestingly, the group of viruses containing NS1 from the allele B clade were ubiquitously overrepresented in all models tested. This is consistent with recent studies suggesting that allele B NS1-containing viruses do not exhibit specific host restrictions or attenuation in replication or in pathogenicity (Turnbull et al., 2016). Finally, we tested the reproducibility of our library system approach both *in vivo* and *in vitro*. To this end, two independent library preparations were compared in infection experiments of murine lungs (Data S1a) and of MDCK cells (Data S1b), showing high correlation profiles at 48 h post-infection (Data S1c).

We next wanted to explore in more detail the relationship between differences in NS1 amino acid sequence versus differences in barcode read abundance among all viruses. For this purpose, we conducted a multiple comparison between differences in NS1 amino acid sequence and differences in fitness between all viruses within the library as described in STAR Methods. Network analysis in murine lungs (Figure 2B), MDCK cells (Figure 2C), and embryonated eggs (Figure 2D) revealed clusters of viruses that shared similar fitness profiles but were significantly different in terms of amino acid sequence. This may indicate a converging evolution for some NS1 gene subsets for host-specific factors. Altogether, we conclude that our barcode-based NGS technique can be efficiently used to track fitness of individual viruses in different hosts within the NS1 viral library population and revealed a large diversity of NS1-driven phenotypic behaviors in multiple host substrates, with evidence for both convergent and divergent evolutions within the NS1 sequence space.

We also analyzed the recombinant library profile using human lung adenocarcinoma A549 cells (Figure S2). Although profile comparison between MDCK and A549 cells displayed differences among several clades of the phylogenetic tree, viruses expressing allele B NS1 were overrepresented within the viral population. Viruses containing H5N1 NS1 sequences were more adapted to MDCK cells than to human A549 cells, while viruses containing NS1 derived from equine and canine hosts were overrepresented after passaging in both cell lines. Despite possible cell line-intrinsic differences, only human A549 cells provided a replicative substrate in which human H3N2 NS1 did not became underrepresented, suggesting that human H3N2 NS1 has become specifically adapted to human cells. This was not the case for human H1N1 NS1.

We next wanted to validate whether the viral library profile dynamics can be reproduced in single-virus experiments using a selection of three viruses within the library with different barcode abundance profiles in all hosts tested (Figure 3A). 10-day-old embryonated chicken eggs (Figure 3B), MDCK cells (Figure 3C), and 8-week-old C57BL/6 mice (Figure 3D) were inoculated with 100 PFU of A/Moscow/2/2007, A/Tasmania/277/2007, or A/Udorn/1972 recombinant viruses, and viral replication was analyzed, as well as body weight loss *in vivo* (Figure 3E). No statistical differences were found in the replication kinetics of the three viruses in eggs and in MDCK cells. A possible reason for the lack of differences in viral titers between these viruses at later time points in these substrates is that they might provide an unlimited resource of replication that only becomes limited when these viruses are put in competition. However, levels of viral replication in mouse lungs were consistent with the data obtained in the context of the library, and this correlated with significant differences in morbidity and mortality among the three viruses tested.

The NS1 protein interacts not only with a multitude of host cell factors but also with other viral proteins to regulate viral gene expression (Marión et al., 1997). To address whether the interactions between our NS1 proteins and other viral proteins of the H1N1 background are contributing to the observed changes within the library, we measured viral replication using recombinant viruses expressing A/Moscow/2/2007, A/Tasmania/277/2007, or A/Udorn/1972 NS1 proteins but in an A/Vietnam/1203/04 (H5N1) HALo background, both *in vitro* and *in vivo* (Figure S3). H5N1 viruses are phylogenetically farther from H1N1 when compared with H3N2 viruses. Thus, if any viral segment other than NS is involved in the phenotypic changes observed, they would become more evident by using an H5N1 viral background. While differences in replication among the



(legend on next page)

three viruses in eggs and in MDCK cells were more apparent than when the PR8 background was used, these differences were consistent with the differences observed in barcode abundance for these NS1 genes when in the PR8 library (Figure 3A). This suggests that the host, not the viral strain background, is the major driver for the fitness differences observed within the library.

Next, we analyzed the multiple NS1 evolutionary directions based on their host-specific fitness within the viral library. To this end, we generated an identity matrix based on amino acid sequence differences and analyzed the $\log_2$ induction over the input at 48 h post-infection for each virus. Data were plotted as three-dimensional (3D) landscapes to visualize the distribution of specific viral groups and their contribution to the overall fitness profile (Figure 4). While allele B NS1-containing viruses shared similar high viral fitness profiles in murine lungs (Figure 4B), MDCK cells (Figure 4D), and allantoic fluid (Figure 4F), allele A H3N2 NS1-expressing viruses showed poor fitness in all three host models tested. However, the distribution of some avian allele A viruses was different among mice (Figure 4A), MDCK cells (Figure 4C), and allantoic fluid (Figure 4E). Among them, recombinant viruses containing H7N9 (A/Shanghai/02/2013) and H9N2 (A/Chicken/Rizhao/2013) NS1 showed increased viral fitness in mouse lungs. Intriguingly, multiple H7N9 human infections have been reported in China since 2013, with more than 1,560 reported human infections as of December 2017 (Shan et al., 2019). Moreover, it has been recently described that H9N2 viruses can be involved in spillover events leading to human infections (Yuan et al., 2017).

To address whether the interferon (IFN) response plays a role in the NS1-dependent fitness profile, *Stat1*$^{-/-}$ mice, which are deficient in IFN signaling, or wild-type (WT) mice (Durbin et al., 2000) were challenged with the viral library. Replication in lungs, body weight loss, and survival data confirmed increased morbidity and mortality in *Stat1*$^{-/-}$ animals infected with the whole virus library (Figure S4). As expected, differences within the viral library were less prominent in infected *Stat1*$^{-/-}$ animals when compared with WT, because several viruses replicated more efficiently in IFN-deficient conditions (Figures 5 and S5A). Such is the case of the recombinant allele B NS1 viruses, for which enrichment was lost in the absence of STAT1, despite still exhibiting increased viral fitness compared with the rest of the library. Viruses containing NS1 sequences from H7N9, H9N2, and H10N4 strains also showed increased fitness levels in *Stat1*$^{-/-}$ mice when compared with their WT counterparts. Thus, we concluded that differences in fitness within these viruses resulted from their NS1-dependent ability to counteract the IFN-mediated antiviral response in WT mice. However, many other viruses in the library, including those with human H3N2 NS1 sequences, were still underrepresented after replication in *Stat1*$^{-/-}$ mice, indicating an NS1-regulated host factor independent of IFN restricts the fitness of these viruses in murine lungs.

In contrast to the results in *Stat1*$^{-/-}$ mice, the behavior for the individual viruses did not change when the library of viruses was used to infect *Rag1*$^{-/-}$ mice, which lack a functional adaptive immune response (Figures S5B and S6). In summary, we conclude that NS1-mediated fitness is largely defined by the early events of the cellular innate immune system.

Finally, we analyzed the temporal fitness dynamics of the viral library in WT mice (Figure 6). Viruses were sorted into different groups following a medoid-based clustering analysis according to their viral fitness properties over time (Figure 6A; Data S2). Viruses that were phylogenetically related shared similar fitness trends. This is exemplified by cluster 4, mainly composed of human H3N2 NS1-containing recombinant viruses, which displayed a decreasing trend from day 2 to day 5 post-infection because of their inability to establish robust levels of viral replication *in vivo* (Data S2d). Cluster 6 contained a heterogenous group of NS1 recombinant viruses (Data S2f). In particular, avian-origin NS1 strains like A/Shanghai/02/2013 (H7N9) and the phylogenetically related A/Chicken/Rizhao/2013 (H9N2), A/Shorebird/DEBay/2009 (H10N7), and A/Mallard/RGeorgia/2011 (H10N4) were phenotypically grouped with all pH1N1 NS1 recombinant viruses included in the library. Despite their phylogenetic differences, all viruses within this cluster shared similar dynamics over time, with increasing overall fitness from day 2 to day 5 post-infection. Finally, allele B NS1-containing viruses were distributed in different clusters, mainly clusters 3 and 7 (Data S2c and S2g). These clusters represent a high and stable fitness profile at all post-infection times tested.

This clustering analysis allowed us to study the combination of viral trends and fitness profiles within the library over time and revealed that the fittest viruses followed two significantly different profiles (Figure 6B). While allele B NS1-containing viruses displayed early replication kinetics that were sustained during the time of the experiment, the best adapted viruses containing allele A NS1 (A/Shanghai/02/2013 and A/Chicken/Rizhao/2013) initially exhibited lower levels of fitness followed by a rapid increase in abundance after day 3 post-infection, reaching comparable levels to the allele B counterparts at day 5 post-infection. We propose that this group of NS1, which includes those of some H7N9 viruses, confers enhanced fitness abilities under high inflammatory conditions that in mice start after day 3 post-infection (Hermesh et al., 2010).

DISCUSSION

Overall, we have used an unbiased system to assess the influenza NS1-mediated tropism in the context of a heterogeneous viral population. Our approach takes advantage of the defined

Figure 1. Phylogenetic and Spatio-temporal Diversity of the NS1-Barcoded Viral Library

Phylogenetic tree containing all available NS1 sequences, assembled by following the neighbor joining (NJ) method and using p distances. Selected NS1 sequences used in this study are depicted as colored arrows along the tree based on their host origin: human (red), avian (blue), swine (green), and other mammals (gray). The distance bar is shown below the tree. Strains are depicted as shown in Table S1. Asterisks denote isolates that were involved in cross-species transmissions following this order: Vietnam/1203/2004 (H5N1) and A/Shanghai/02/2013 (H7N9) viruses were included as avian origin (blue arrows), because they came from a zoonotic transmission from an avian host. Similarly, A/Elephant seal/California/2010 is red (human), because it was a reverse zoonotic transmission from humans.

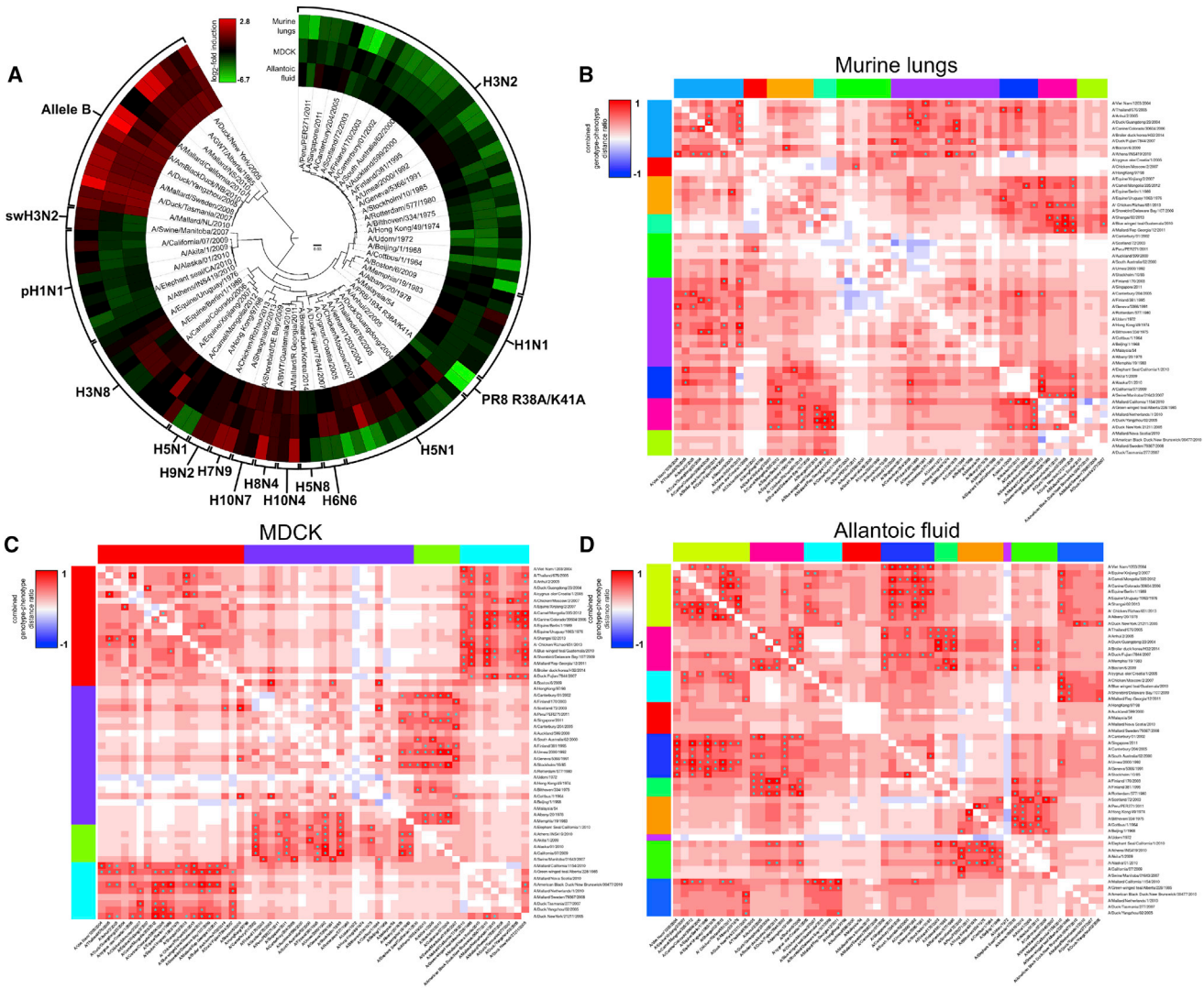


Figure 2. Analysis of the Population Dynamics within the NS1-Barcoded Influenza Viral Library Reveals Over- and Underrepresented Subgroups upon Infection *In Vivo*, *In Vitro*, and *In Ovo*

(A) Circular heatmap displaying the relative abundance of barcode reads for each recombinant virus within the library upon infection of 10-day-old embryonated chicken eggs, MDCK cells, or 8-week-old C57BL/6 mice. Samples were collected at 48 h post-infection (hpi), and after viral RNA extraction, they were further processed to quantify the total number of barcode reads. Averages of triplicates are listed in columns around the phylogenetic tree and expressed as the $\log_2$ fold induction over the relative proportion of barcode reads found in the initial viral mix (input). Red and green indicate high or low barcode reads versus input, respectively.

(B–D) Multiple comparison between differences in amino acid sequence versus differences in fitness between all viruses within the library were analyzed in mouse lungs (B), in MDCK supernatants (C), and in allantoic fluid (D) at 48 hpi. Highly correlating viruses were grouped in clusters labeled in different colors for visualization purposes. Pairwise comparisons showing viruses with similar fitness profiles but with significant amino acid differences are depicted in red. *$p < 0.05$.

sequence space within naturally occurring IAV strains. Using a library of IAVs containing different NS1 representatives of the known NS1 genotypic variation, we found multiple evolutionary lines for this viral protein, leading to a great variety of host tropism phenotypes. We found phenotypic differences between NS1 phenotypes that were evolutionary related, such as the ones of human H3N2 viruses and old seasonal H1N1 viruses. This is unexpected, because H3N2 and H1N1 NS1 share a common H1N1 ancestor (the NS1 of the 1918 IAV) and both have been circulating in humans for decades without reassorting since 1958 (Medina and García-Sastre, 2011). The different phenotypes of these two groups of human NS1 sequences are reminiscent of speciation of one species into two when separated by geographical constraints and highlight the constant evolution of NS1.

We also found that sustained circulation of NS1 in a particular host may or may not result in specialization. Such is the case of viruses containing H3N2 NS1, which showed preference to human hosts. In contrast, we found that viruses with avian-origin allele B NS1 were not restricted to a particular

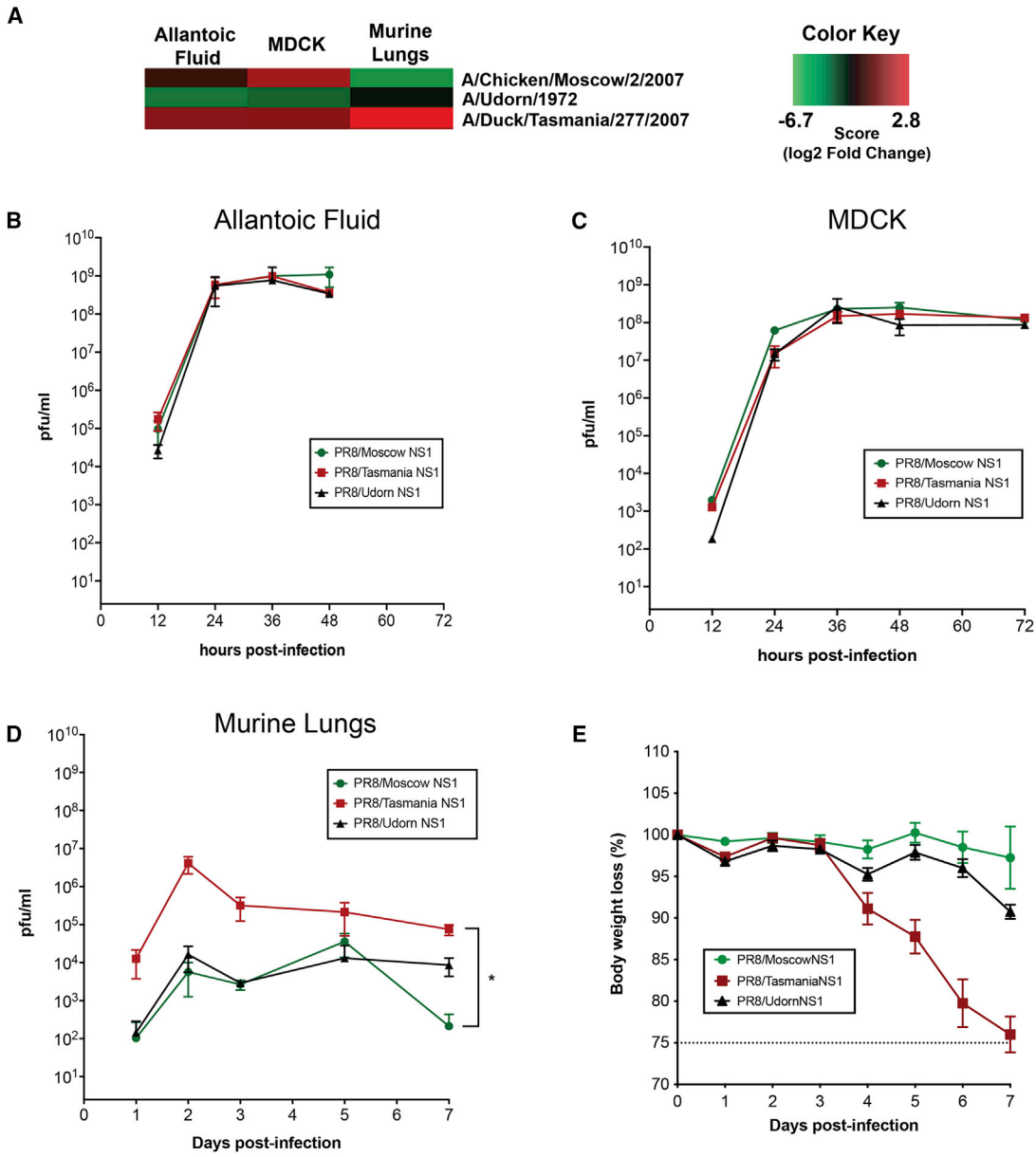


Figure 3. Viral Library Profile Dynamics Can Be Reproduced in Single-Virus Experiments in a A/Puerto Rico/8/1934 (H1N1) Background *In Vivo*

(A) Three representative viruses showing different viral trends in the library were selected based on differences in barcode abundance.

(B–D) Single-virus infections were conducted in triplicate using 10-day-old embryonated chicken eggs (B), MDCK cells (C), and 8-week-old C57BL/6 mice (D). Viral replication was quantified at different time points post-infection.

(E) Body weight loss of infected mice was monitored daily.

Error bars depict the SD. *$p < 0.05$.

host, because they had efficient replication profiles in all systems tested. While it has previously been shown that allele B-derived segments efficiently replicated in mammalian hosts despite their avian origin (Turnbull et al., 2016), the specific contributions of NS1 and NEP were not explored. In our current study, we addressed the replication abilities of a curated selection of NS1 proteins in a NEP-independent manner because of our split NS construct strategy. We can conclude that allele B NS1 is able to provide efficient replication fitness in a range of hosts.

In our study, we cannot distinguish between fitness effects due to the species origin and those due to the substrate origin of the virus growth substrates. In other words, while differences in viral fitness between MDCK and A549 cells are likely related to their species of origin (canine versus human), they might also be influenced by the tissue origin (MDCK cells are kidney cells, and

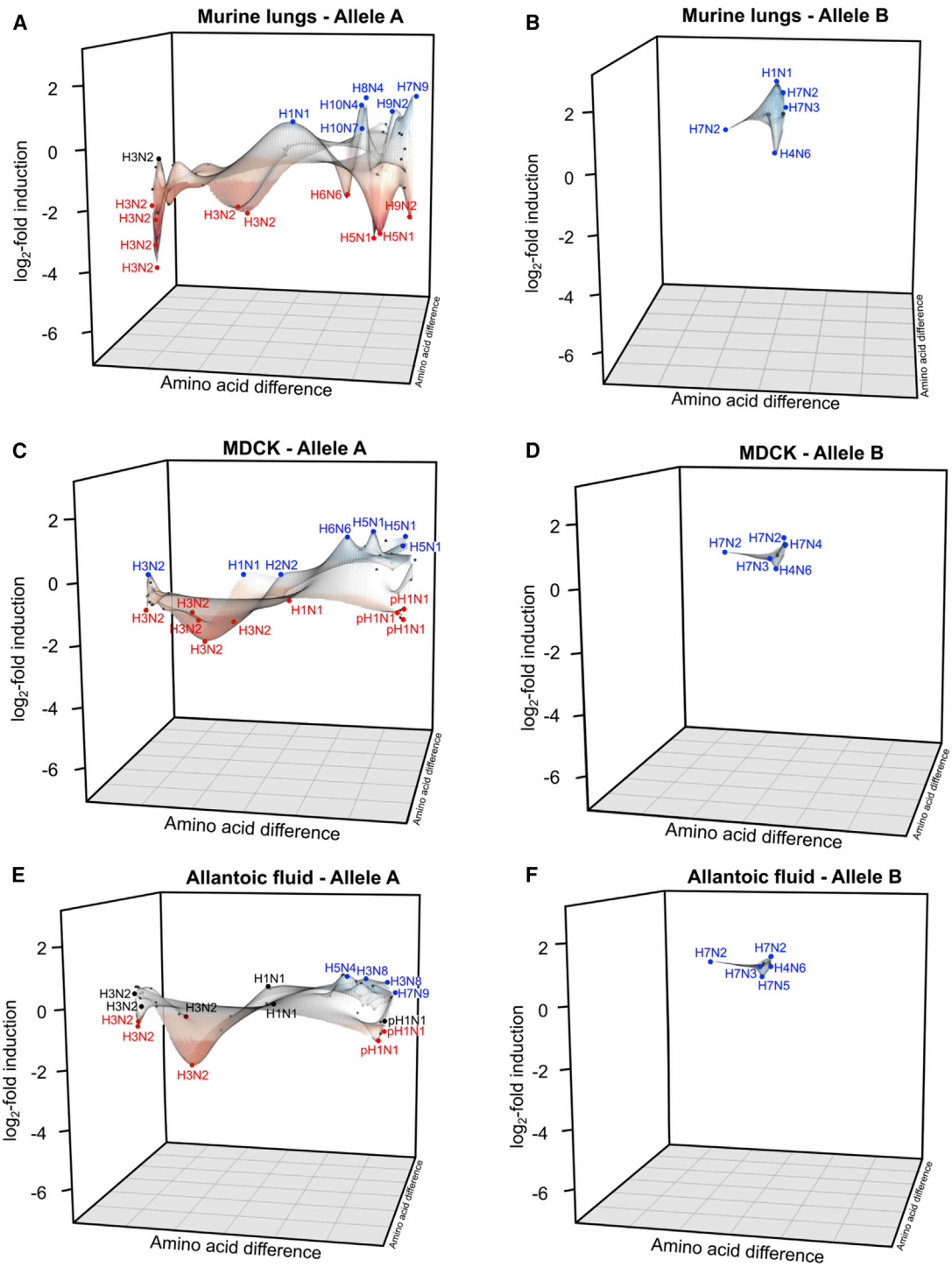


Figure 4. NS1 Evolutionary Directions within the Viral Library Show Species-Dependent Profiles that Preferentially Overrepresent Specific NS1 Clades

Amino acid differences across all NS1 sequences were integrated in a percentage identity matrix, and relative distances were plotted using x and y coordinates. Averages of triplicates expressed as the $\log_2$ fold induction over the initial relative proportion of barcode reads found in the initial viral mix (input) were plotted on the z axis. (A–F) Individual plots in mice (A and B), MDCK cells (C and D), and allantoic fluid (E and F) were generated. Viruses containing allele A (A, C, E) or allele B (B, D, F) NS1 were plotted separately in each case for visualization purposes.

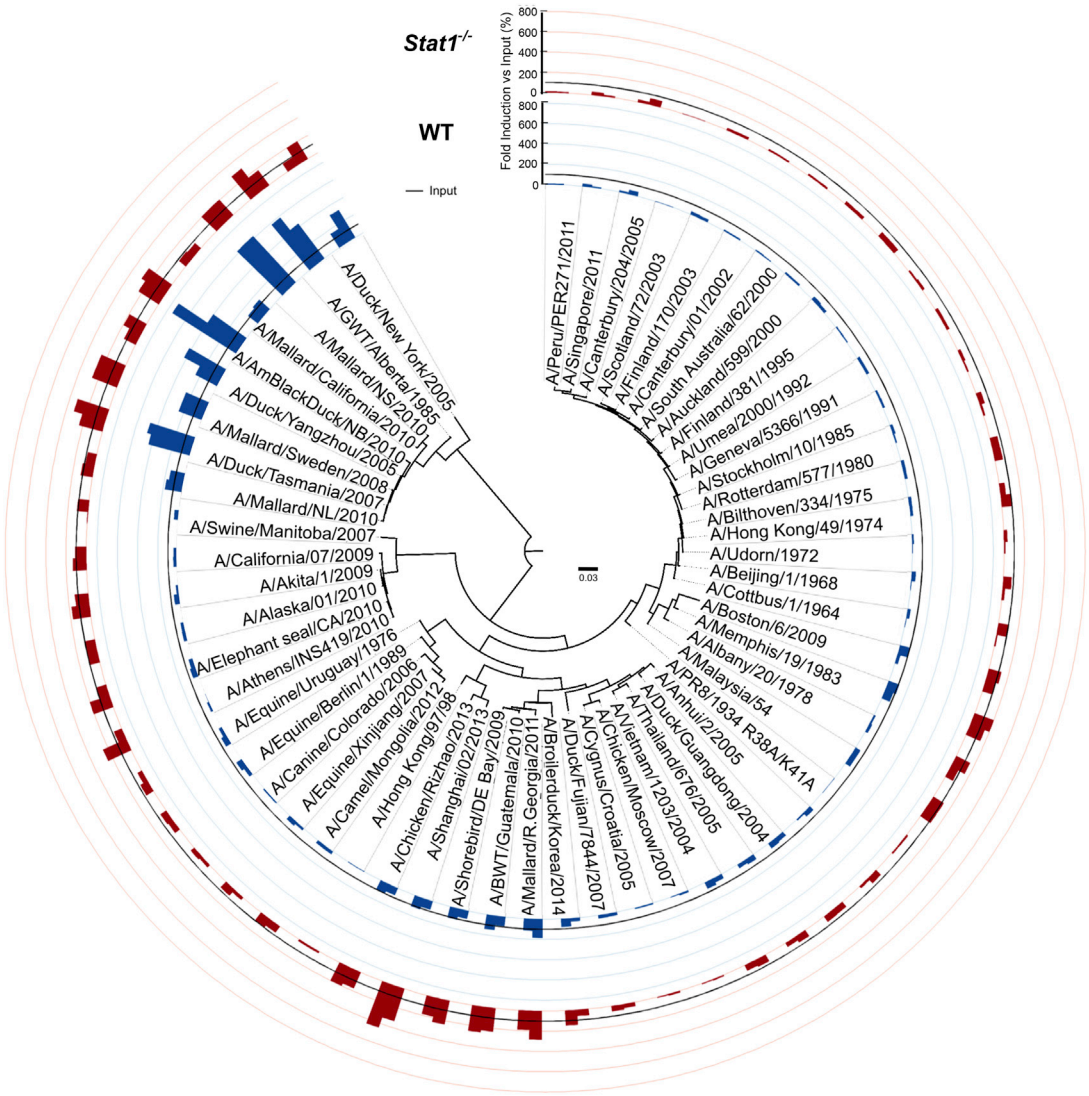
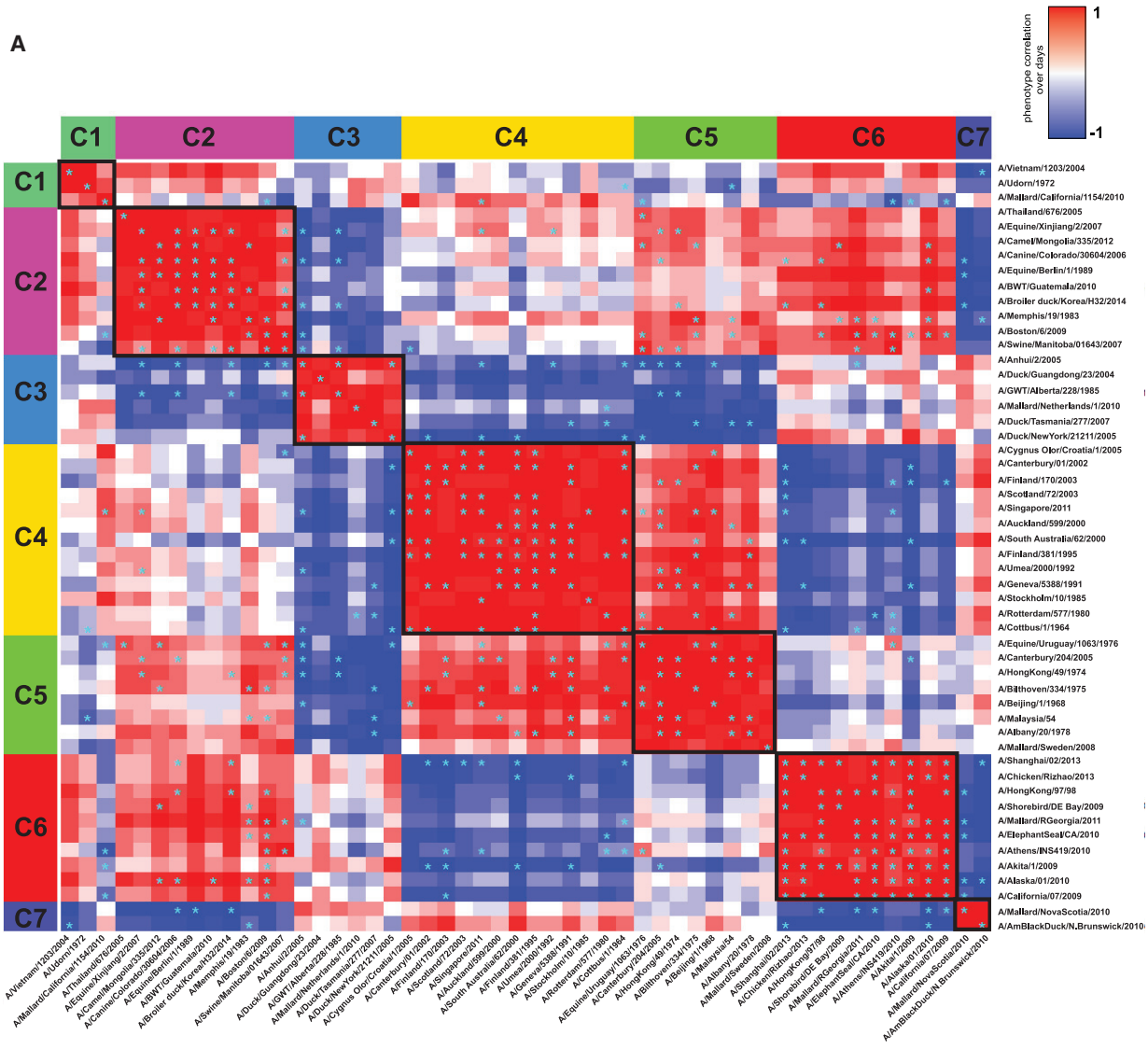


Figure 5. NS1 Viral Library Fitness Profile Is Innate Immune Response Dependent

Circular bar graph comparing results obtained from infected WT and *Stat1*$^{-/-}$ mice at day 3 post-infection in triplicate and expressed as the relative barcode fold increase percentage over the input.

A549 are respiratory epithelial cells) or even by the passage history of the cell line. Nevertheless, differences between these substrates reflect a differential impact of host factors and tissue environment in NS1-mediated fitness.

Thus, we propose that differences in tropism among the IAV NS1 genes result from adaptations to specific host factors among the large variety of known NS1 interactors. For instance, NS1 proteins that interact with highly conserved host proteins to enhance viral replication might display a broad host tropism. Such is the case of the host polyadenylation factor CPSF30, which is one of the known host targets interacting with NS1 (Nemeroff et al., 1998). In contrast, NS1 variants that do not interact with CPSF30, like the 2009 pandemic H1N1 virus, may have adapted to inhibit other host factors with high sequence diversity among the host species, like TRIM25. This would result in host specialization of this subgroup of NS1. Nevertheless, focusing on a particular domain or specific amino acid changes within the NS1 sequence may not be the right approach to understand the virus-host interactions and the mechanisms behind the multiple phenotypes observed in our library experiments. We found evidence of such complexity in our results with a subset of NS1 that confers increased fitness under high pro-inflammatory conditions *in vivo*, as well as fitness differences in a type I IFN-dependent manner. Our future work will rely on multiscale, high-throughput studies that will explore protein-protein interactions between NS1 and cellular host factors. This will provide further insights into how the virus modulates the antiviral response and how it evolves through this dynamic selective pressure.



(legend on next page)

In summary, our barcoded library approach can aid in the identification of IAV strains with increased host tropism and thus higher pandemic risk potential. Our system can be applied to viruses with high sequence diversity besides IAVs, facilitating analysis of specific viral genes and their contribution to the overall viral fitness in different hosts and/or cell types.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines
  - Mouse lines
  - Chicken embryonated eggs
- METHOD DETAILS
  - Selection of NS1 sequences to include in the library
  - IAV infections
  - Deep sequencing and data processing of the NS1 recombinant library
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - IAV infections
  - Networks and time series heatmaps
  - Phylogenetic analysis
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.11.070.

ACKNOWLEDGMENTS

We thank Dr. Kimihito Ito (Hokkaido University) for sharing the R code that was used in the evolutionary direction studies. We thank Mrs. Magdalena Aichinger for her support and technical help throughout the generation and edition of the 3D landscape figures, as well as Mr. Richard Cadagan and Mr. Osman Lizardo (Icahn School of Medicine at Mount Sinai) for providing technical assistance. D.B.-M. is an Open Philanthropy Fellow of the Life Sciences Research Foundation. The figures of this work were designed to facilitate their interpretation to readers with color vision deficiencies. However, due to the complexity of specific data graphs, some panels might require color conversion software available for free, such as Visolve (https://www.ryobi-sol.co.jp/visolve/en/). This research work was supported partially by CRIP (Center for Research on Influenza Pathogenesis), an NIH-NIAID-funded Center of Excellence for Influenza Research and Surveillance (CEIRS) (contract HHSN272201400008C), and by NIH-NIAID (grants U19AI117873 and U19AI135972). This research was also supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF), funded by the Ministry of Science and ICT of the Republic of Korea (NRF-2018M3A9H4056537 to M.-S.P., PI).

AUTHOR CONTRIBUTIONS

Conceptualization, R.M.-M., C.M.-R., and A.G.-S.; Methodology, R.M.-M., C.M.-R., D.B.-M., C.V.F., R.N., A.A.B., I.M., S.A., V.B., I.L., M.P., J.A., and D.S.; Investigation, R.M.-M. and C.M.-R.; Writing – Original Draft, R.M.-M., C.M.-R., and A.G.-S.; Writing – Review & Editing, R.M.-M., C.M.-R., A.G.-S., and B.T.; Funding Acquisition, A.G.-S. and M.S.-P.; Resources, A.G.-S., M.S.-P., and F.K.; Supervision, A.G.-S. and B.R.t.O.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 15, 2019
Revised: July 3, 2019
Accepted: November 15, 2019
Published: December 17, 2019

REFERENCES

Ayllon, J., and García-Sastre, A. (2015). The NS1 protein: a multitasking virulence factor. Curr. Top. Microbiol. Immunol. *386*, 73–107.

Benitez, A.A., Panis, M., Xue, J., Varble, A., Shim, J.V., Frick, A.L., López, C.B., Sachs, D., and tenOever, B.R. (2015). *In Vivo* RNAi Screening Identifies MDA5 as a Significant Contributor to the Cellular Defense against Influenza A Virus. Cell Rep. *11*, 1714–1726.

Bergmann, M., Garcia-Sastre, A., Carnero, E., Pehamberger, H., Wolff, K., Palese, P., and Muster, T. (2000). Influenza virus NS1 protein counteracts PKR-mediated inhibition of replication. J. Virol. *74*, 6203–6206.

Bodenhofer, U., Bonatesta, E., Horejš-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. Bioinformatics *31*, 3997–3999.

Borg, I., and Groenen, P.J.F. (2005). Modern Multidimensional Scaling: Theory and Applications (Springer).

Donelan, N.R., Basler, C.F., and García-Sastre, A. (2003). A recombinant influenza A virus expressing an RNA-binding-defective NS1 protein induces high levels of beta interferon and is attenuated in mice. J. Virol. *77*, 13257–13266.

Durbin, J.E., Fernandez-Sesma, A., Lee, C.K., Rao, T.D., Frey, A.B., Moran, T.M., Vukmanovic, S., García-Sastre, A., and Levy, D.E. (2000). Type I IFN modulates innate and specific antiviral immunity. J. Immunol. *164*, 4220–4228.

Fodor, E., Devenish, L., Engelhardt, O.G., Palese, P., Brownlee, G.G., and García-Sastre, A. (1999). Rescue of influenza A virus from recombinant DNA. J. Virol. *73*, 9679–9682.

Forst, C.V. (2000). Molecular evolution of catalysis. J. Theor. Biol. *205*, 409–431.

García-Sastre, A., Egorov, A., Matassov, D., Brandt, S., Levy, D.E., Durbin, J.E., Palese, P., and Muster, T. (1998). Influenza A virus lacking the NS1 gene replicates in interferon-deficient systems. Virology *252*, 324–330.

Greenspan, D., Palese, P., and Krystal, M. (1988). Two nuclear location signals in the influenza virus NS1 nonstructural protein. J. Virol. *62*, 3020–3026.

Guo, Z., Chen, L.M., Zeng, H., Gomez, J.A., Plowden, J., Fujita, T., Katz, J.M., Donis, R.O., and Sambhara, S. (2007). NS1 protein of influenza A virus inhibits the function of intracytoplasmic pathogen sensor, RIG-I. Am. J. Respir. Cell Mol. Biol. *36*, 263–269.

Hale, B.G., Steel, J., Medina, R.A., Manicassamy, B., Ye, J., Hickman, D., Hai, R., Schmolke, M., Lowen, A.C., Perez, D.R., and García-Sastre, A. (2010).

Figure 6. Time Course Analysis Revealed Correlating Clusters of Phylogenetically Related Viruses with Similar Trends *In Vivo*

(A) WT mice were infected with the NS1 viral library, and barcode read profiles for each virus were analyzed at days 2–5 post-infection. Fitness trends over time were clustered by similarity. *$p < 0.05$.

(B) Representative viruses exhibiting the highest barcode abundance followed different dynamics over time and can be mainly divided into two groups. For instance, allele B NS1 recombinant viruses showed high levels since early after infection (continuous lines), while others, such as A/Shanghai/02/2013 (H7N9) and its phylogenetically closest related A/Chicken/Rizhao/651/2013 (H9N2), gradually increased over time. Error bars depict SD of at least 6 replicates.

Inefficient control of host gene expression by the 2009 pandemic H1N1 influenza A virus NS1 protein. J. Virol. *84*, 6909–6922.

Hatada, E., and Fukuda, R. (1992). Binding of influenza A virus NS1 protein to dsRNA *in vitro*. J. Gen. Virol. *73*, 3325–3329.

Hermesh, T., Moltedo, B., López, C.B., and Moran, T.M. (2010). Buying time-the immune system determinants of the incubation period to respiratory viruses. Viruses *2*, 2541–2558.

Ito, K., Igarashi, M., Miyazaki, Y., Murakami, T., Iida, S., Kida, H., and Takada, A. (2011). Gnarled-trunk evolutionary model of influenza A virus hemagglutinin. PLoS ONE *6*, e25953.

Kerr, M.K. (2003). Linear models for microarray data analysis: hidden similarities and differences. J. Comput. Biol. *10*, 891–901.

Kochs, G., García-Sastre, A., and Martínez-Sobrido, L. (2007). Multiple anti-interferon actions of the influenza A virus NS1 protein. J. Virol. *81*, 7011–7021.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645.

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol. Biol. Evol. *33*, 1870–1874.

Li, S., Min, J.Y., Krug, R.M., and Sen, G.C. (2006). Binding of the influenza A virus NS1 protein to PKR mediates the inhibition of its activation by either PACT or double-stranded RNA. Virology *349*, 13–21.

Liu, J., Lynch, P.A., Chien, C.Y., Montelione, G.T., Krug, R.M., and Berman, H.M. (1997). Crystal structure of the unique RNA-binding domain of the influenza virus NS1 protein. Nat. Struct. Biol. *4*, 896–899.

Ludwig, S., Schultz, U., Mandler, J., Fitch, W.M., and Scholtissek, C. (1991). Phylogenetic relationship of the nonstructural (NS) genes of influenza A viruses. Virology *183*, 566–577.

Marión, R.M., Zürcher, T., de la Luna, S., and Ortín, J. (1997). Influenza virus NS1 protein interacts with viral transcription-replication complexes *in vivo*. J. Gen. Virol. *78*, 2447–2451.

McKenzie, A.T., Katsyv, I., Song, W.M., Wang, M., and Zhang, B. (2016). DGCA: A comprehensive R package for Differential Gene Correlation Analysis. BMC Syst. Biol. *10*, 106.

Medina, R.A., and García-Sastre, A. (2011). Influenza A viruses: new research developments. Nat. Rev. Microbiol. *9*, 590–603.

Min, J.Y., and Krug, R.M. (2006). The primary function of RNA binding by the influenza A virus NS1 protein in infected cells: Inhibiting the 2′-5′ oligo (A) synthetase/RNase L pathway. Proc. Natl. Acad. Sci. USA *103*, 7100–7105.

Munir, M., Zohari, S., Metreveli, G., Baule, C., Belák, S., and Berg, M. (2011). Alleles A and B of non-structural protein 1 of avian influenza A viruses differentially inhibit beta interferon production in human and mink lung cells. J. Gen. Virol. *92*, 2111–2121.

Nachbagauer, R., Choi, A., Hirsh, A., Margine, I., Iida, S., Barrera, A., Ferres, M., Albrecht, R.A., García-Sastre, A., Bouvier, N.M., et al. (2017). Defining the antibody cross-reactome directed against the influenza virus surface glycoproteins. Nat. Immunol. *18*, 464–473.

Nemeroff, M.E., Barabino, S.M., Li, Y., Keller, W., and Krug, R.M. (1998). Influenza virus NS1 protein interacts with the cellular 30 kDa subunit of CPSF and inhibits 3'end formation of cellular pre-mRNAs. Mol. Cell *1*, 991–1000.

Neumann, G., Watanabe, T., Ito, H., Watanabe, S., Goto, H., Gao, P., Hughes, M., Perez, D.R., Donis, R., Hoffmann, E., et al. (1999). Generation of influenza A viruses entirely from cloned cDNAs. Proc. Natl. Acad. Sci. USA *96*, 9345–9350.

Noronha, J.M., Liu, M., Squires, R.B., Pickett, B.E., Hale, B.G., Air, G.M., Galloway, S.E., Takimoto, T., Schmolke, M., Hunt, V., et al. (2012). Influenza virus sequence feature variant type analysis: evidence of a role for NS1 in influenza virus host range restriction. J. Virol. *86*, 5857–5866.

Palese, P., and Shaw, M.L. (2007). Orthomyxoviridae: the viruses and their replication. In Fields Virology, D.M. Knipe and P.M. Howley, eds. (Lippincott Williams & Wilkins), pp. 1647–1689.

Reynolds, A.P., Richards, G., de la Iglesia, B., and Rayward-Smith, V.J. (2006). Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms. JMMA *5*, 475–504.

Satterly, N., Tsai, P.L., van Deursen, J., Nussenzveig, D.R., Wang, Y., Faria, P.A., Levay, A., Levy, D.E., and Fontoura, B.M. (2007). Influenza virus targets the mRNA export machinery and the nuclear pore complex. Proc. Natl. Acad. Sci. USA *104*, 1853–1858.

Shan, X., Lai, S., Liao, H., Li, Z., Lan, Y., and Yang, W. (2019). The epidemic potential of avian influenza A (H7N9) virus in humans in mainland China: A two-stage risk analysis. PLoS ONE *14*, e0215857.

Talon, J., Horvath, C.M., Polley, R., Basler, C.F., Muster, T., Palese, P., and García-Sastre, A. (2000). Activation of interferon regulatory factor 3 is inhibited by the influenza A virus NS1 protein. J. Virol. *74*, 7989–7996.

Treanor, J. (2004). Influenza vaccine—outmaneuvering antigenic shift and drift. N. Engl. J. Med. *350*, 218–220.

Turnbull, M.L., Wise, H.M., Nicol, M.Q., Smith, N., Dunfee, R.L., Beard, P.M., Jagger, B.W., Ligertwood, Y., Hardisty, G.R., Xiao, H., et al. (2016). Role of the B Allele of Influenza A Virus Segment 8 in Setting Mammalian Host Range and Pathogenicity. J. Virol. *90*, 9263–9284.

Varble, A., Albrecht, R.A., Backes, S., Crumiller, M., Bouvier, N.M., Sachs, D., García-Sastre, A., and tenOever, B.R. (2014). Influenza A virus transmission bottlenecks are defined by infection route and recipient host. Cell Host Microbe *16*, 691–700.

Yuan, R., Liang, L., Wu, J., Kang, Y., Song, Y., Zou, L., Zhang, X., Ni, H., and Ke, C. (2017). Human infection with an avian influenza A/H9N2 virus in Guangdong in 2016. J. Infect. *74*, 422–425.

Zohari, S., Munir, M., Metreveli, G., Belák, S., and Berg, M. (2010). Differences in the ability to suppress interferon β production between allele A and allele B NS1 proteins from H10 influenza A viruses. Virol. J. *7*, 376.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
Influenza A virus (A/Puerto Rico/8/34/Mount Sinai(H1N1))	This manuscript	NCBI: txid183764
Influenza A virus (A/Vietnam/1203/2004(H5N1))	This manuscript	NCBI: txid284218
Chemicals, Peptides, and Recombinant Proteins		
Lipofectamine 2000	Thermo Fisher Scientific	Cat #: 11668-019
Purified Agar	Oxoid	Cat #: LP0028
SuperScript III One-Step RT-PCR System	Thermo Fisher Scientific	Cat #: 12574018
Superscript III Reverse Transcriptase	Thermo Fisher Scientific	Cat #: 18080044
Trypsin from bovine pancreas, TPCK-treated	SIGMA-ALDRICH	Cat #: T1426-500MG
Critical Commercial Assays		
E.Z.N.A Gel Extraction Kit (V-spin)	Omega Bio-Tek	Cat #: D2500-01
KAPA Library Quantification Kit - Illumina GA	Roche	Cat #: KK4824
MiSeq Reagent Kit v3	Illumina	Cat #: MS-102-3001
QIAamp Viral RNA Mini Kit (250)	QIAGEN	Cat #: 52906
TruSeq DNA Library Preparation Kit v2, Set A	Illumina	Cat #: RS-122-2001
TruSeq RNA Sample Prep Kit v2 -Set B	Illumina	Cat #: RS-122-2002
Deposited Data		
Illumina deep-sequencing raw data	This manuscript	https://data.mendeley.com/datasets/77x6pyfk79/1 https://data.mendeley.com/datasets/ryr859j6pr/1
Experimental Models: Cell Lines		
Dog: MDCK cells	ATCC	Cat #: CCL-34; RRID: CVCL_0422
Human: 293T cells	ATCC	Cat #: CRL-3216; RRID: CVCL_0063
Human: A549 cells	ATCC	Cat #: CCL-185; RRID: CVCL_0023
Experimental Models: Organisms/Strains		
Mouse: B6.129S7-Rag1tm1Mom/J	Jackson Laboratories	Cat #: 002216; RRID: IMSR_JAX:002216
Mouse: C57BL/6J	Jackson Laboratories	Cat #: 000664; RRID: IMSR_JAX:000664
Mouse: 129S6/SvEv-Stat1tm1Rds	Taconic	Cat #: 2045-F; RRID: IMSR_TAC:2045
Mouse: 129S6/SvEvTac	Taconic	Cat #: 129SVE-F; RRID: IMSR_TAC:129sve
Chicken: Specific Pathogen Free Fertile Eggs	Charles River	Cat #: 10100329
Oligonucleotides		
Barcodes used in the viral library	This manuscript	See Table S1
Primers used in the Illumina deep-sequencing analysis	This manuscript	See Table S2
Recombinant DNA		
NS1 sequences used in the library	This manuscript	See Table 1
Software and Algorithms		
Barcode Generator 2.8	The Comai Lab	http://comailab.genomecenter.ucdavis.edu/index.php/Barcode_generator
Circos 0.69-4	Krzywinski et al., 2009	http://circos.ca/software/download/circos/
FigTree 1.4.4	The Rambaut Lab	http://tree.bio.ed.ac.uk/software/figtree/
Geneious 9.1.5	Geneious	https://www.geneious.com/
Illumina Basespace	Illumina	https://login.illumina.com/platform-services-manager/?rURL=https://basespace.illumina.com&clientId=basespace&clientVars=aHR0cHM6Ly9iYXNlc3BhY2UuaWxsdW1pbmEuY29tL2Rhc2hib2FyZA&redirectMethod=GET#/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MATLAB 9.7	MathWorks	https://www.mathworks.com/products/matlab.html
MEGA7	Kumar et al., 2016	https://www.megasoftware.net/
MeV	GitHub	https://sourceforge.net/projects/mev-tm4/
Prism 8.2.1	Graphpad	https://www.graphpad.com
RStudio 1.2.5001	Rstudio, Inc.	https://rstudio.com
Other		
Customs R scripts to analyze specific datasets of this work	This manuscript	https://data.mendeley.com/datasets/ryr859j6pr/1
R package cluster	CRAN	https://cran.r-project.org/web/packages/cluster/
R package DGCA	CRAN	https://cran.r-project.org/web/packages/DGCA/
R package fpc	CRAN	https://cran.r-project.org/web/packages/fpc/
R package ggplot2	CRAN	https://cran.r-project.org/web/packages/ggplot2/
R package gplots	CRAN	https://cran.r-project.org/web/packages/gplots/
R package igraph	CRAN	https://cran.r-project.org/web/packages/igraph/
R package limma	CRAN	https://cran.r-project.org/src/contrib/Archive/limma/
R package MBA	CRAN	https://cran.r-project.org/web/packages/MBA/
R package msa	BIOINF	http://www.bioinf.jku.at/software/msa/
R package RColorBrewer	CRAN	https://cran.r-project.org/web/packages/RColorBrewer/
R package rgl	CRAN	https://cran.r-project.org/web/packages/rgl/
R package seqinr	CRAN	https://cran.r-project.org/web/packages/seqinr/
R package sna	CRAN	https://cran.r-project.org/web/packages/sna/
R package splines	CRAN	https://cran.r-project.org/src/contrib/Archive/splines/
R package xlsx	CRAN	https://cran.r-project.org/web/packages/xlsx/

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for reagents may be directed to and will be fulfilled by the Lead Contact, Adolfo García-Sastre (adolfo.garcia-sastre@mssm.edu). All stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

The female origin HEK293T, male origin A549 and female origin MDCK cell lines were obtained from the American Type Culture Collection (ATCC) Manassas, VA) and cultured in Dulbecco's modified Eagle's medium (Thermo Fisher Scientific, Waltham, MA) supplemented with 10% fetal bovine serum (FBS) (Thermo Fisher Scientific, Waltham, MA), 100 IU/ml Penicillin, 100 μg/ml Streptomycin and 0.25 μg/ml Amphotericin B (Corning Antibiotic-Antimycotic Solution) (CORNING, Corning, NY), at 37°C in a 5% CO_2 incubator.

Mouse lines

The mouse line *B6.129S7-Rag1*^{tm1Mom}/J (referred to as *Rag1*$^{-/-}$) and control animals of the C57BL/6J strain were obtained from The Jackson Laboratory (Bar Harbor, ME). The mouse line 129S6/SvEv-*Stat1*^{tm1Rds} (referred to as *Stat1*$^{-/-}$) and control animals of the 129S6/SvEvTac strain (referred as 129S6) were obtained from Taconic (Rensselaer, NY). For all animal experiments, 6- to 8-week-old female mice were randomly assigned to experimental groups and housed under Specific Pathogen Free (SPF) conditions. All animal studies were reviewed and approved by the Institutional Animal Care and Use Committee (IACUC) of the Icahn School of Medicine at Mount Sinai, in accordance with the institutional and national guidelines and regulations and performed in Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC)-certified facilities.

Chicken embryonated eggs

10-day-old specific-pathogen-free (SPF) chicken embryonated eggs were obtained from Charles River Laboratories (Wilmington, MA) and incubated at 37°C and 55%–60% humidity conditions.

METHOD DETAILS

Selection of NS1 sequences to include in the library

NS1 sequences from different virus isolates were selected from the Influenza Research Database (IRD) based on different structural and functional Sequence Feature Variant Types (SFVT) (Noronha et al., 2012). To obtain a robust diversity within the different NS1 variants, two different strains were selected for each sequence feature (SF): the first included in variant 1 (VT1), that uses A/Udorn/1972 as strain reference, and the second one corresponded to a different VT containing a relevant number of amino acid variation for that specific SF. The strain selection was refined after multiple alignments, discarding NS1 sequences that were identical and only considering the strains whose genome was fully sequenced and represented multiple times in the database. In addition, other parameters, such as the country of origin, host and year were also considered. Following this approach, a list of the most representative NS1 was created. As an internal control, each NS1 recombinant virus was rescued in duplicates and associated to a different and unique 22-nucleotide barcode. Barcodes were designed using a specific software (Barcode generator) and they were all GC content matched as previously reported (Varble et al., 2014). A total of 107 viruses out of 112 were rescued, plaque purified, titered and combined in equal amounts (100 pfu per virus) to constitute the final library (Table 1). Only few NS1 recombinant viruses were more difficult to rescue, thus resulting in ~4% of viruses in the library that are not paired to a second barcode. This was likely due to variable rescue efficiencies when different plasmid preparations were used. Nevertheless, these few viruses represented by a single barcode were included as part of the final library since no major differences in barcode relative abundance were detected between viral pairs containing different barcodes. Viruses represented only with a single barcode are A/Auckland/599/2000, A/Malaysia/1954, A/Hong Kong/97/98, A/Mallard/Nova Scotia/00088/2010 and A/Mallard/Sweden/79367/2008. Barcode read ratios obtained for each recombinant virus within the original library mix (input) were close to 1:1:1, thus minimizing any amplification bias within the library construction. A small library containing 4 recombinant viruses (A/Shanghai/02/2013, A/Udorn/1972, A/Puerto Rico/8/1934 and A/Puerto Rico/8/1934-R38A/K41A) was used as a proof of principle to our strategy (Figure S1), before assembling the final viral library.

Plasmid-based viral rescue and propagation

A modified NS segment that introduces a non-coding region between NS1 and NEP was used as previously described (Varble et al., 2014). Each NS segment was individually designed with a unique 22-nucleotide neutral barcode, inserted in the non-coding region between NS1 and NEP. This labeling strategy allows to individually track and quantify each NS segment in the Illumina Deep Sequencing platform. A total of 56 NS1 sequences were selected for the final library. As an internal control, two different barcodes were generally used per each NS1 sequence. NS1 sequences were synthesized using the GeneArt Strings DNA fragments platform (Thermo Fisher Scientific, Waltham, MA) and cloned into EcoRI and NheI sites into the pDZ vector. Standard reverse genetics were used to rescue each individual virus in a A/Puerto Rico/8/1934 (PR8) backbone, as previously described (Fodor et al., 1999). Lipofectamine 2000 (Thermo Fisher Scientific, Waltham, MA) was used to transfect a co-culture of HEK293T and MDCK cells with seven ambisense pDZ plasmids in combination with the pDZ NS1 split barcoded plasmid specific for each IAV, and incubated for 72 hours at 37°C in a 5% CO_2 incubator. Additionally, plasmid-based rescue system for recombinant viruses in the A/Vietnam/1203/04/H5N1 (HALo) backbone followed a slightly different strategy, using 8 pPolI plasmids encoding the viral RNAs and 4 pCAGGS expression plasmids encoding the viral proteins PB1, PB2, PA and NP. Cell culture supernatants containing influenza virions were collected. Final stocks of each recombinant virus were achieved after plaque purification and infecting fresh MDCK cells (80% confluence) with a 1:100 dilution of the purified plaque and incubated at 37°C for 48 hours in Minimum Essential Medium Eagle (MEM) (Thermo Fisher Scientific, Waltham, MA) supplemented with 1μg/ml of TPCK trypsin (Sigma-Aldrich, St. Louis, MO). Viral titers were determined by plaque assay and by hemagglutination assays using a suspension of 5% turkey red blood cells in PBS. Absence of defective interfering particles was confirmed by comparing hemagglutination activity versus viral titers.

IAV infections

Viral infections were performed either at a specific multiplicity of infection (MOI) or by combining 100 pfu per virus when using the viral library. Cell lines were infected with the viral inoculum in PBS supplemented with 0.3% Bovine Serum Albumin (BSA) (Gemini, West Sacramento, CA) in triplicates. After one hour of incubation at room temperature, cells were washed with DMEM supplemented with 10% FBS to remove the unbound virus. Cells were incubated in MEM supplemented with 0.3% BSA and either with 0.2 μg/ml (A549 cells) or 1 μg/ml (MDCK cells) of TPCK-trypsin at 37°C in a 5% CO_2 incubator. When indicated, cell culture supernatants were sampled at different time points to assess viral replication by standard plaque assay technique using MDCK cells. 10-day-old embryonated eggs were infected with 100 μl of viral library in PBS in triplicates. 48 hours later, allantoic fluid was harvested for downstream processing. Infection experiments in all mouse models tested were done in triplicate by inoculating each animal intranasally with 25 μl of PBS containing either the viral library or individual recombinant viruses. Animals were monitored daily for weight loss and clinical signs of disease. At the indicated times post infection, animals were euthanized and lung tissue was collected in 1 mL of PBS. Samples were homogenized using sterile ceramic beads after two rounds of mechanical treatment of 10 s at 6.5 m/s. Tissue debris was removed by three rounds of low-speed centrifugations and viral titers in the supernatants were determined by plaque assay in triplicates. Samples were further processed by NGS techniques to measure relative viral barcode abundance.

Deep sequencing and data processing of the NS1 recombinant library

Viral RNA extraction from supernatants of infected MDCK or A549 cells, allantoic fluid and mouse lungs was performed using the QIAamp Viral RNA Mini Kit (QIAGEN, Valencia, CA), following manufacturer's instructions. In order to monitor viral populations, Superscript III One step RT-PCR (Thermo Fisher Scientific, Waltham, MA) was used with specific primers to amplify the whole NS segment. This was followed by a nested PCR strategy to amplify the region surrounding the barcode sequence (~300 bp), using specific primers that contained barcoded Illumina linkers. This step was performed using a low cycle count to reduce possible nucleotide mismatches during the amplification process. Up to twenty samples were multiplexed per run and then analyzed on the Illumina MiSeq sequencing platform. Barcoded reads were obtained and analyzed with a custom-specific software consisting of R-written scripts and run on Mount Sinai's high-performance computing cluster following the parameters previously described (Varble et al., 2014). Over time propagation of viral populations was visualized using MATLAB and/or Prism 8 software. Additionally, all NS1 recombinant viruses in the library were full genome deep-sequenced. For that, previously extracted viral RNA were heat fragmented followed by first and second strand cDNA synthesis using the Superscript III Reverse Transcriptase (Thermo Fisher Scientific, Waltham, MA). Then, the 3′ end of each sample was repaired and further ligated with adapters for multiplex purposes. Samples were enriched by PCR and bands with fragment size between 200 bp and 400 bp were cut and purified from an agarose gel for further processing. Finally, DNA amount for each sample was quantified by RT-PCR and viral samples were combined, denatured, mixed with the internal control PhiX and loaded into the Illumina cartridge for sequencing. Circular plots were built using Circos software (Krzywinski et al., 2009).

NS1 tridimensional landscapes

Three dimensional landscapes were generated in a similar manner as in a previous publication (Nachbagauer et al., 2017). To generate the horizontal plane for the 3D landscapes, amino acid sequences of all NS1 proteins were aligned with the Clustal Omega Multiple Sequence Alignment tool and a Percent Identity Matrix was generated. The distances were calculated by multi-dimensional scaling based on the percent amino acid sequence difference (Borg and Groenen, 2005; Ito et al., 2011). The x and y coordinates therefore represent the relative distance between NS1 proteins based on percent amino acid difference. The z axis indicates the relative up- or downregulation on a $\log_2$ scale. The NS1 sequences were separated into two groups (alleles A and B) for presentation. The scales of all axes are consistent between the graphs. The surfaces were approximated from the $\log_2$ upregulation or downregulation data for all NS1 proteins using multilevel B-splines. The ***mba.surf*** function implemented in the Multilevel B-spline Approximation package in R version 3.2.5 was used.

Networks and time series heatmaps

All distance calculations were done using custom R scripts. Amino acid sequences of the different NS1 proteins were first aligned using the multiple sequence alignment (***msa***) R package (Bodenhofer et al., 2015) and ClustalW with a BloSum62 similarity matrix and standard parameters. Pairwise distances were then converted into a distance matrix for further processing and phenotypic data was $\log_2$ transformed. Organism-individual as well as combined datasets from the individual measurements were further processed. Data from single case/control measurements were converted into distances by Euclidian metric. Time series data was transformed into distances d using both Euclidian metric and Pearson's correlation derived distance $d\rho = (1-\rho)/2$. Genotypic and phenotypic distances were normalized to assure $0 \leq d \leq 1$ and further combined by linear combination assuming identical weights ($= 1$) for both genotypic and phenotypic contributions. Combined distances were then converted into similarities by $sim = d_{max} - d$. Differential distances between pairs of genotypes and phenotypes were calculated by using the Differential Gene Correlation Analysis (***DGCA***) R package (McKenzie et al., 2016). Normalized genotype and phenotype distances were calculated as previously described (Forst, 2000). Distances were first converted to similarities for further processing together with corresponding sample size information. Then, resulting matrices of scaled differences in distances were further clustered. Cluster analysis of correlation and distance data as well as results from differential distance analysis was performed by using a *Partitioning Around Medoids* (Reynolds et al., 2006). The number of clusters (k) was estimated by optimum average silhouette width using the ***fpc*** R package. Similarity graphs were visualized with CytoScape using a force directed layout with weights based on the similarity measure. Similarity clusters have been identified by hierarchical clustering based on the combined distance or by calculating the optimized community structure based on the weighted modularity from the corresponding weighted similarity graph. Outliers between virus duplicates were assessed by a hierarchical linear model using ***Limma*** (Kerr, 2003) with time and replicate as two factors.

QUANTIFICATION AND STATISTICAL ANALYSIS

IAV infections

A Kruskal-Wallis test was used to compare the differences between single-virus experiments, followed by Dunn's multiple comparisons post-test correction. Student's t test was used to analyze abundance profile changes between WT, *Stat1*$^{-/-}$ and *Rag1*$^{-/-}$ mice, followed by Welch's correction post-test. Two-way ANOVA was used to validate the reproducibility of two sets of viral libraries, followed by Bonferroni's multiple comparison post-test correction. The data were processed with Microsoft Excel and/or Apple Numbers, and plotted with GraphPad Prism 8.

Networks and time series heatmaps

Similarities between trajectories were assessed by Pearson's correlation together with raw and corrected *p-value*s. Significantly different distances were identified by *p-value*s after ***DGCA***. Significantly different strains that were identified between the duplicated barcode measurements ($p < 0.13$) were considered outliers and excluded for further cluster analysis. Strains with a single barcode measurement were included on the analysis. All *p-value*s were corrected for multiple testing after Benjamini-Hochberg using a cutoff of 0.05 for significance.

Phylogenetic analysis

NS1 protein sequences were aligned using MUSCLE algorithm in MEGA7 (Kumar et al., 2016). Neighbor Joining (NJ) clustering analysis was performed to create a phylogenetic tree with all NS1 sequences available in the databases and further map our NS1 sequences within the library. Bayesian evolutionary analysis (BEAST) was used to create a phylogenetic tree containing our final selection of NS1 sequences within the library. Phylogenetic trees were assembled with Geneious software (version 9.1.5). Further tree visualizations were performed using FigTree software.

DATA AND CODE AVAILABILITY

The datasets and code generated during this study are available at Mendeley Data (https://data.mendeley.com/datasets/77x6pyfk79/1 and https://data.mendeley.com/datasets/ryr859j6pr/1).

Cell Reports, Volume *29*

Supplemental Information

Viral Fitness Landscapes in Diverse Host Species Reveal Multiple Evolutionary Lines for the NS1 Gene of Influenza A Viruses

Raquel Muñoz-Moreno, Carles Martínez-Romero, Daniel Blanco-Melo, Christian V. Forst, Raffael Nachbagauer, Asiel Arturo Benitez, Ignacio Mena, Sadaf Aslam, Vinod Balasubramaniam, Ilseob Lee, Maryline Panis, Juan Ayllón, David Sachs, Man-Seong Park, Florian Krammer, Benjamin R. tenOever, and Adolfo García-Sastre

FIGURE S1

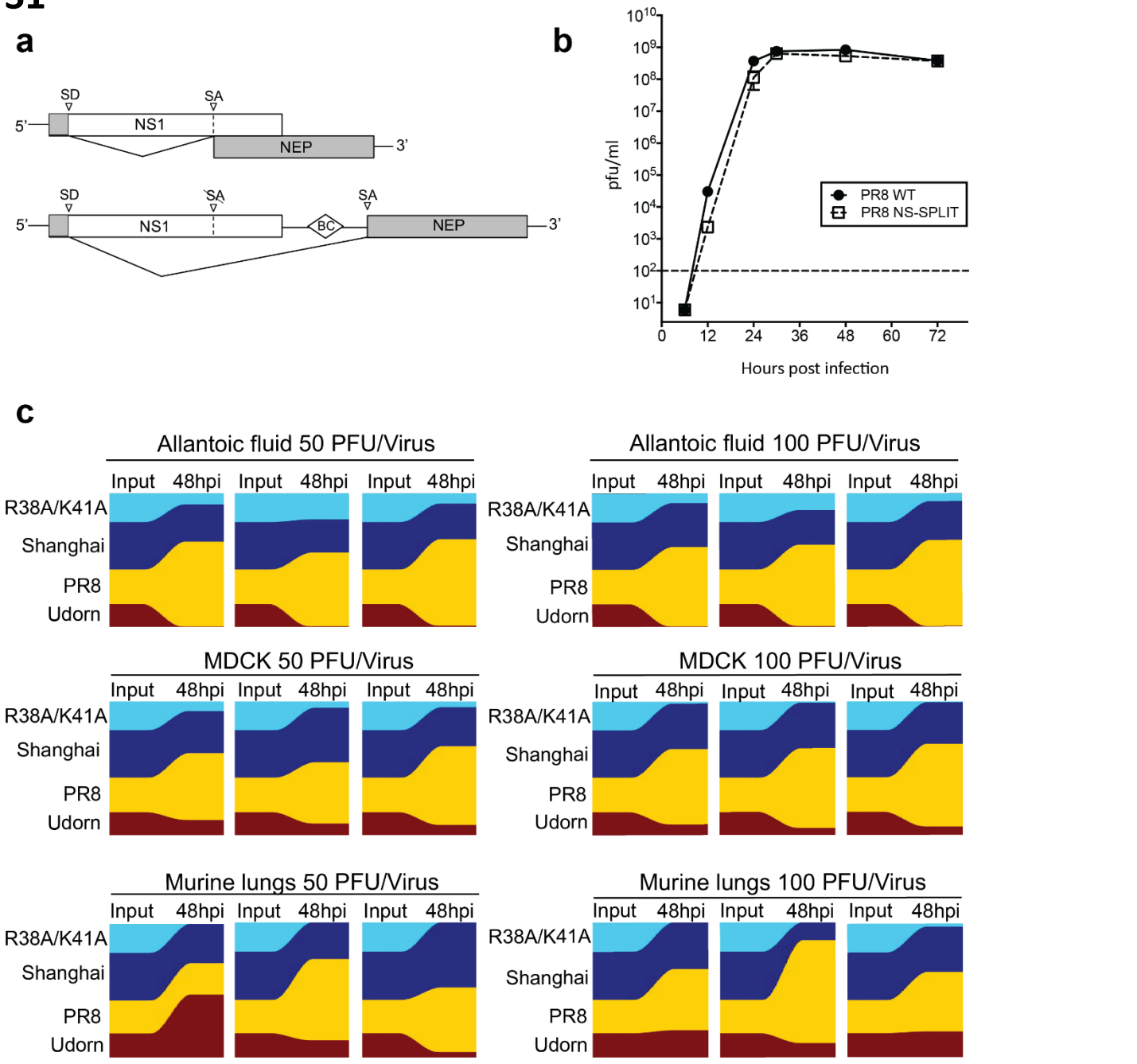


Figure S1, related to Figure 1. Reproducible barcode-based abundance is achieved after infecting *in vitro*, *in ovo* and *in vivo* systems with a small NS1 recombinant library. (a) Schematic representation of WT NS segment (top) and split NS segment (bottom) with modified splicing mechanism that allows the expression of NS1 and NEP in different open reading frames, flanking a short neutral barcode. (**b**) Replication comparison of a WT PR8 virus and a recombinant NS-split PR8 virus. MDCK cells were infected at a MOI of 0.001 pfu/cell. Error bars depict the standard deviation (SD) of three independent experiments. (**c**) Viral progression diagrams (Matlab) depicting the relative viral abundance present in a "proof-of-principle" library relative to the total number of barcode reads. Each color represents the relative proportion of a specific barcode within the viral population. "Input" indicates the proportion within the initial inoculum at the time of infection. 10-day-old chicken embryonated eggs (upper panels), MDCK cells (middle panels) and 8-week-old WT C57BL/6 mice (lower panels) were infected in triplicates with the viral library by adding either 50 pfu/virus or 100 pfu/virus and samples were collected 48 hours post infection.

FIGURE S2

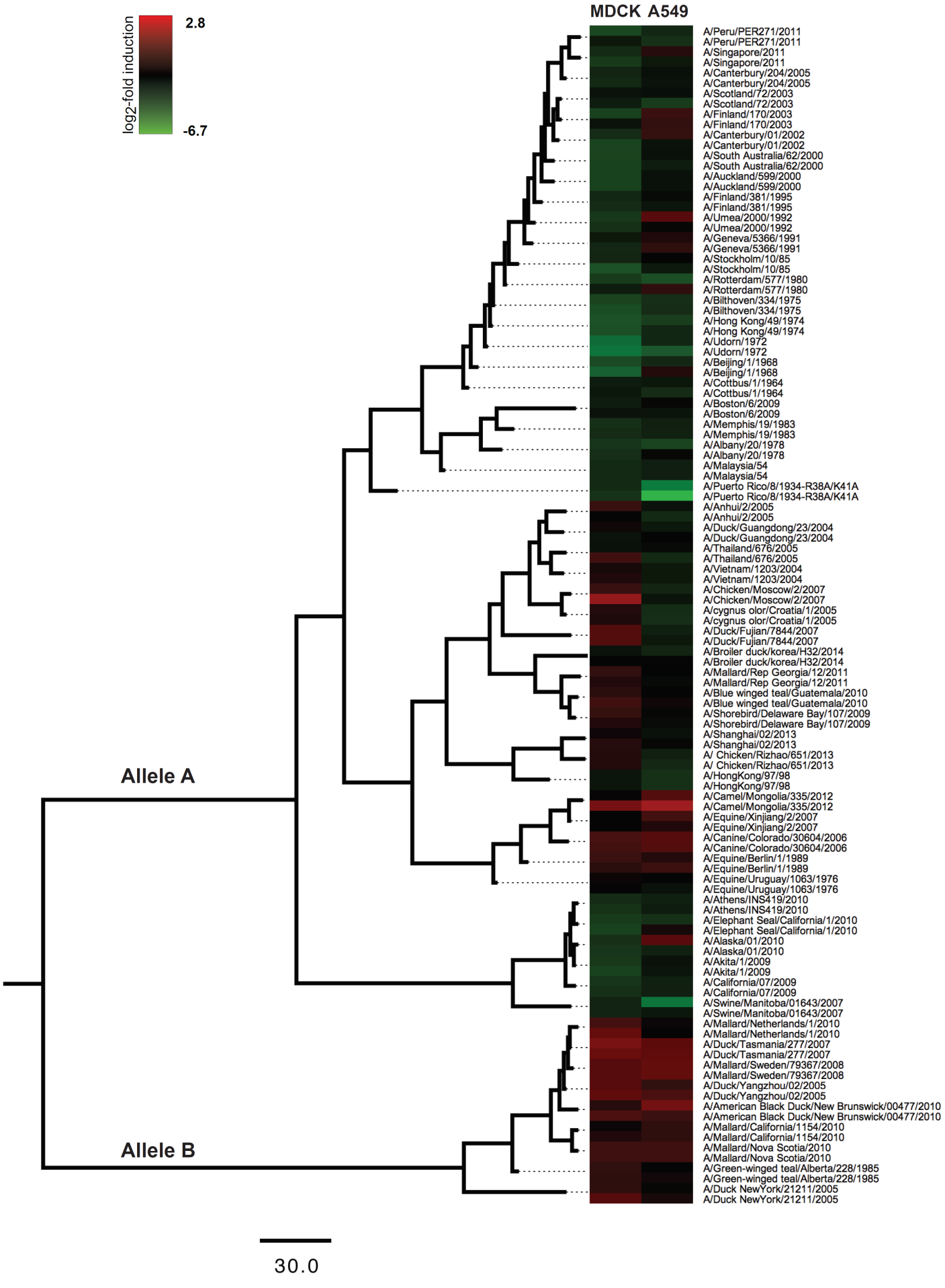

log2-fold induction
2.8
-6.7
MDCK A549
A/Peru/PER271/2011
A/Peru/PER271/2011
A/Singapore/2011
A/Singapore/2011
A/Canterbury/204/2005
A/Canterbury/204/2005
A/Scotland/72/2003
A/Scotland/72/2003
A/Finland/170/2003
A/Finland/170/2003
A/Canterbury/01/2002
A/Canterbury/01/2002
A/South Australia/62/2000
A/South Australia/62/2000
A/Auckland/599/2000
A/Auckland/599/2000
A/Finland/381/1995
A/Finland/381/1995
A/Umea/2000/1992
A/Umea/2000/1992
A/Geneva/5366/1991
A/Geneva/5366/1991
A/Stockholm/10/85
A/Stockholm/10/85
A/Rotterdam/577/1980
A/Rotterdam/577/1980
A/Bilthoven/334/1975
A/Bilthoven/334/1975
A/Hong Kong/49/1974
A/Hong Kong/49/1974
A/Udorn/1972
A/Udorn/1972
A/Beijing/1/1968
A/Beijing/1/1968
A/Cottbus/1/1964
A/Cottbus/1/1964
A/Boston/6/2009
A/Boston/6/2009
A/Memphis/19/1983
A/Memphis/19/1983
A/Albany/20/1978
A/Albany/20/1978
A/Malaysia/54
A/Malaysia/54
A/Puerto Rico/8/1934-R38A/K41A
A/Puerto Rico/8/1934-R38A/K41A
A/Anhui/2/2005
A/Anhui/2/2005
A/Duck/Guangdong/23/2004
A/Duck/Guangdong/23/2004
A/Thailand/676/2005
A/Thailand/676/2005
A/Vietnam/1203/2004
A/Vietnam/1203/2004
A/Chicken/Moscow/2/2007
A/Chicken/Moscow/2/2007
A/cygnus olor/Croatia/1/2005
A/cygnus olor/Croatia/1/2005
A/Duck/Fujian/7844/2007
A/Duck/Fujian/7844/2007
A/Broiler duck/korea/H32/2014
A/Broiler duck/korea/H32/2014
A/Mallard/Rep Georgia/12/2011
A/Mallard/Rep Georgia/12/2011
A/Blue winged teal/Guatemala/2010
A/Blue winged teal/Guatemala/2010
A/Shorebird/Delaware Bay/107/2009
A/Shorebird/Delaware Bay/107/2009
A/Shanghai/02/2013
A/Shanghai/02/2013
A/ Chicken/Rizhao/651/2013
A/ Chicken/Rizhao/651/2013
A/HongKong/97/98
A/HongKong/97/98
A/Camel/Mongolia/335/2012
A/Camel/Mongolia/335/2012
A/Equine/Xinjiang/2/2007
A/Equine/Xinjiang/2/2007
A/Canine/Colorado/30604/2006
A/Canine/Colorado/30604/2006
A/Equine/Berlin/1/1989
A/Equine/Berlin/1/1989
A/Equine/Uruguay/1063/1976
A/Equine/Uruguay/1063/1976
A/Athens/INS419/2010
A/Athens/INS419/2010
A/Elephant Seal/California/1/2010
A/Elephant Seal/California/1/2010
A/Alaska/01/2010
A/Alaska/01/2010
A/Akita/1/2009
A/Akita/1/2009
A/California/07/2009
A/California/07/2009
A/Swine/Manitoba/01643/2007
A/Swine/Manitoba/01643/2007
A/Mallard/Netherlands/1/2010
A/Mallard/Netherlands/1/2010
A/Duck/Tasmania/277/2007
A/Duck/Tasmania/277/2007
A/Mallard/Sweden/79367/2008
A/Mallard/Sweden/79367/2008
A/Duck/Yangzhou/02/2005
A/Duck/Yangzhou/02/2005
A/American Black Duck/New Brunswick/00477/2010
A/American Black Duck/New Brunswick/00477/2010
A/Mallard/California/1154/2010
A/Mallard/California/1154/2010
A/Mallard/Nova Scotia/2010
A/Mallard/Nova Scotia/2010
A/Green-winged teal/Alberta/228/1985
A/Green-winged teal/Alberta/228/1985
A/Duck NewYork/21211/2005
A/Duck NewYork/21211/2005
Allele A
Allele B
30.0

Figure S2, related to Figure 2. Different NS1 selection profiles upon MDCK or A549 cells library infection. Phylogenetic tree containing selected NS1 sequences within the library was assembled following Bayesian analysis (BEAST). Heat map displaying the relative abundance of barcode reads for each recombinant virus within the library upon infection of MDCK or A549 cells. Viral RNA samples were isolated at 48 hours post infection and further analyzed. Average of the triplicates are listed in columns and expressed as the log_2-fold induction over the initial relative proportion of barcode reads found in the initial viral mix (input). Red and green colors indicate high or low barcode representation versus the input, respectively. Additional information of each specific NS1 is available in **Table 1**.

FIGURE S3

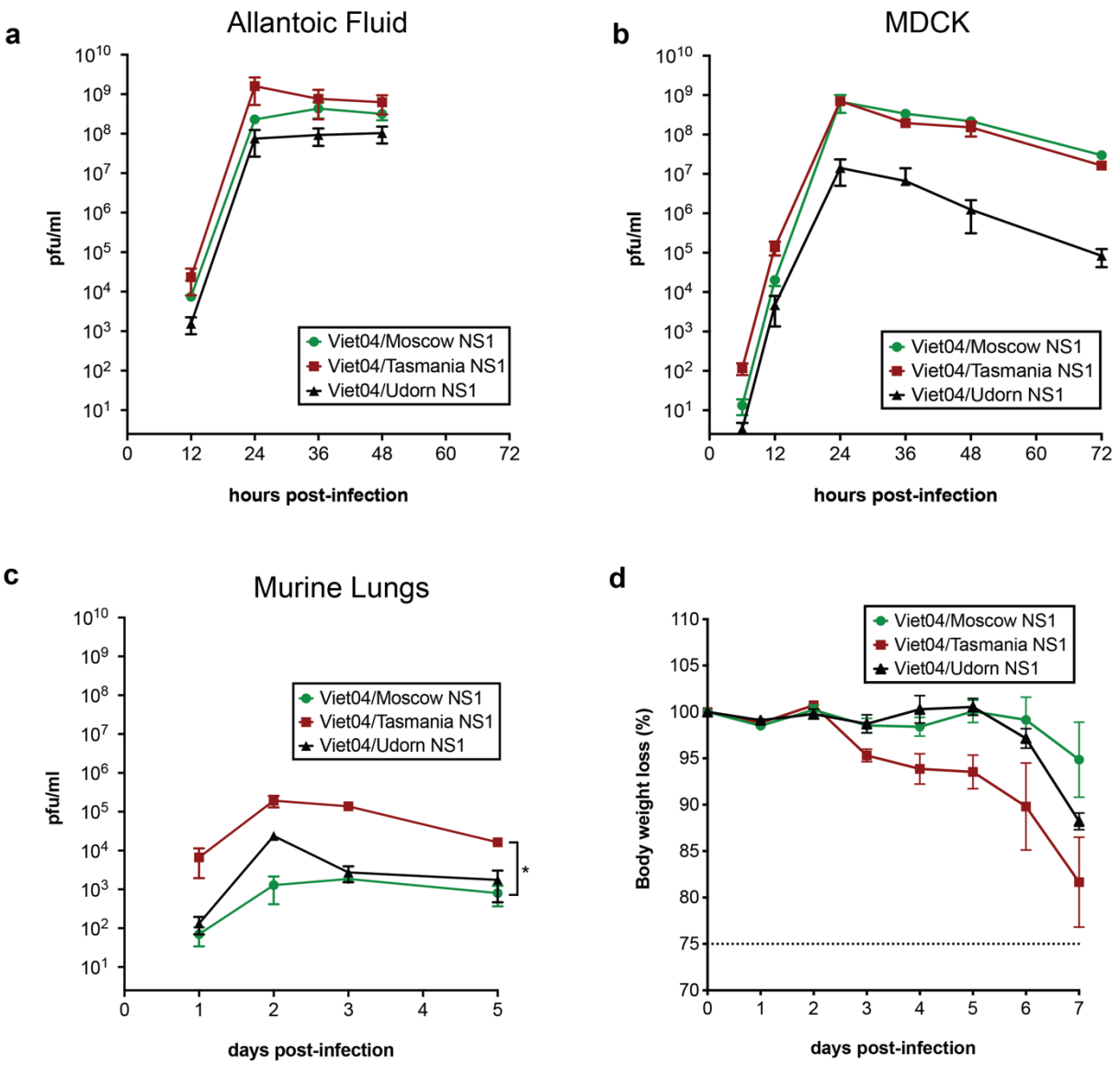


Figure S3, related to Figure 3. Viral library profile dynamics can be reproduced in single-virus experiments in A/Vietnam/1203/04 (H5N1 HALo) background. Same NS1 variants as in Figure 3 were rescued in a H5N1 HALo background and single-virus infections were conducted in triplicates using 10-day-old chicken embryonated eggs (**a**) MDCK cells (**b**) and 8-week-old C57BL/6 mice (**c**). Viral replication was quantified at different time points post infection. Body weight loss of infected mice was daily monitored (**d**). Error bars depict the standard deviation (SD). *, $p<0.05$.

FIGURE S4

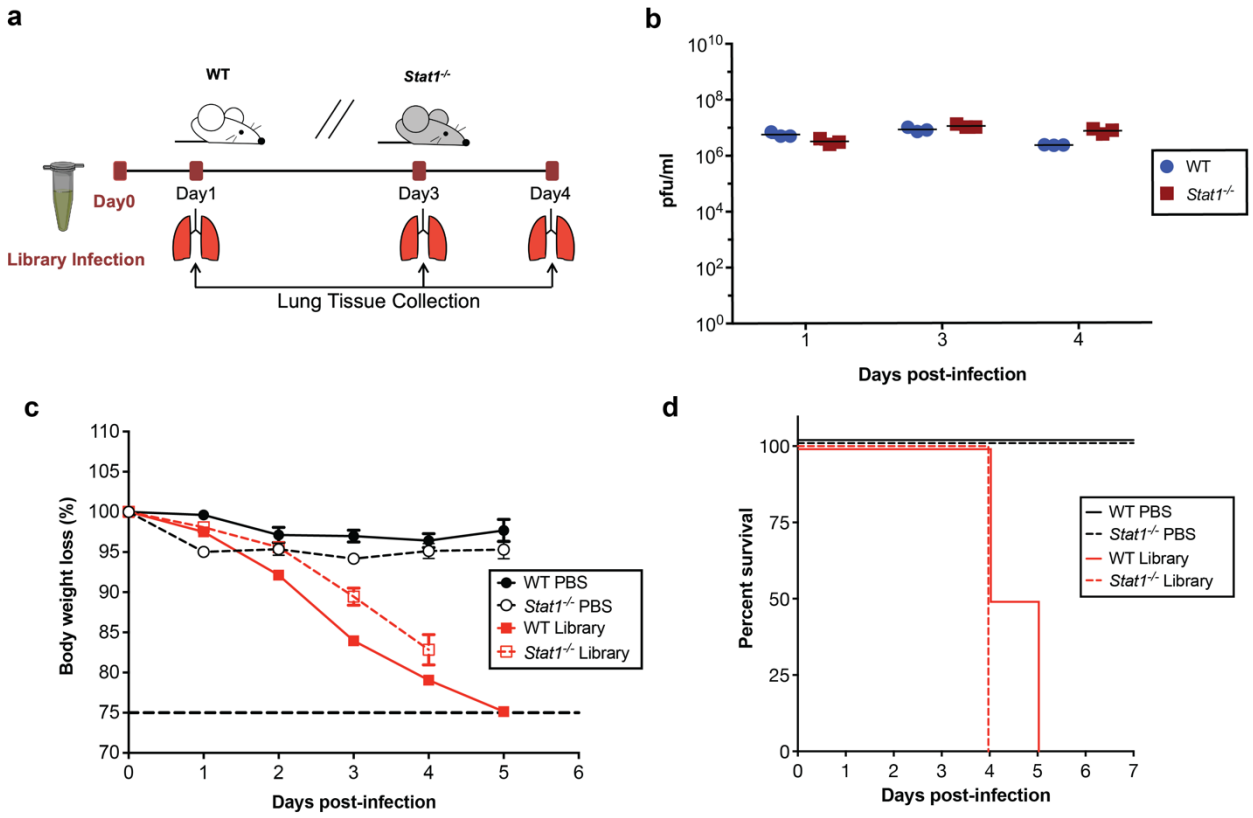


Figure S4, related to Figure 5. NS1 library selection profile depends on the type-I IFN response. (**a**) Experimental layout of library infection in 129S wild type (WT) and 129S *Stat1*$^{-/-}$ mice (n=4). (**b**) Mouse lung viral titers were determined by plaque assay on MDCK cells at days 1, 3 and 4 post-infection. (**c**) Body weight loss and (**d**) survival rates were daily monitored during the infection. Error bars depict the standard deviation (SD).

FIGURE S5

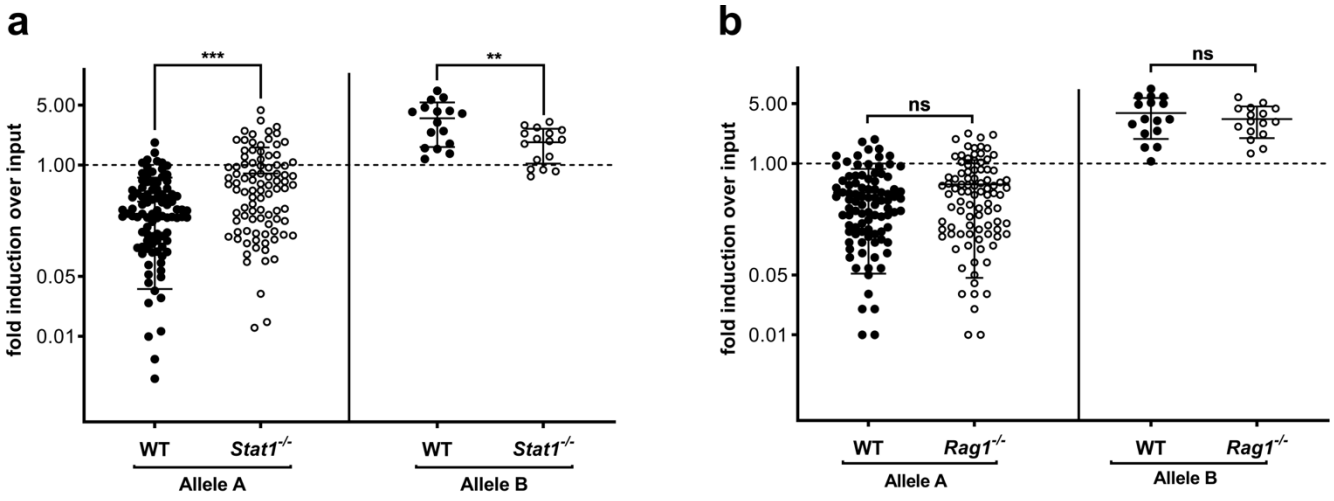


Figure S5, related to Figure 5. Differences in fitness between allele A and allele B NS1. Scatter plot graphs depicting the abundance profiles of viruses containing allele A or allele B NS1 within the library. Average of the triplicates are represented as the fold induction over the input (initial relative proportion of barcode reads found in the viral mix) and represented in a $\log_2$ scale. Profiles from *Stat1*$^{-/-}$ at 3 days post-infection (**a**) and *Rag1*$^{-/-}$ mice at 4 days post infection (**b**) are shown. Error bars depict the standard deviation (SD). *, P<0.05; **, P<0.01; ***, P<0.001.

FIGURE S6

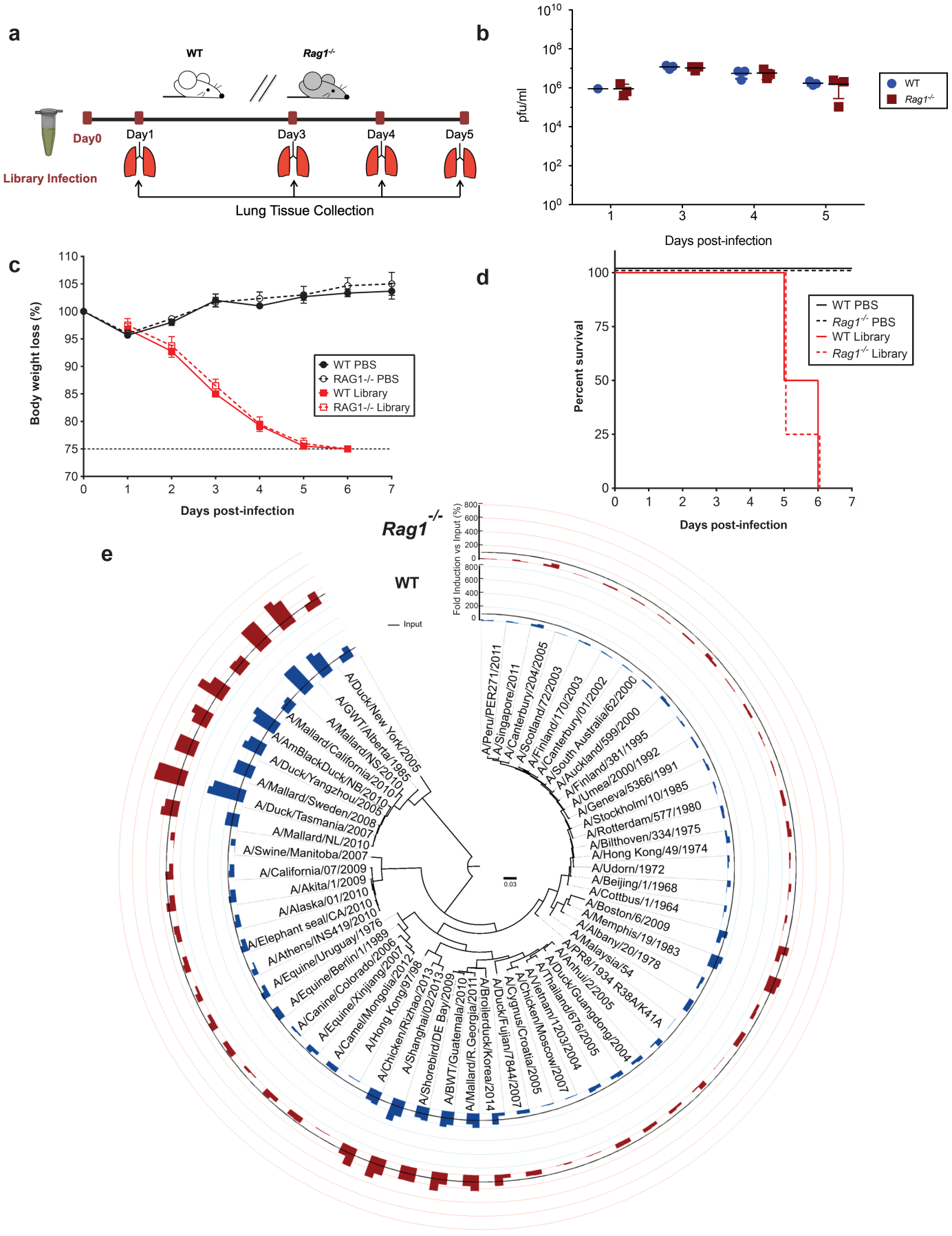
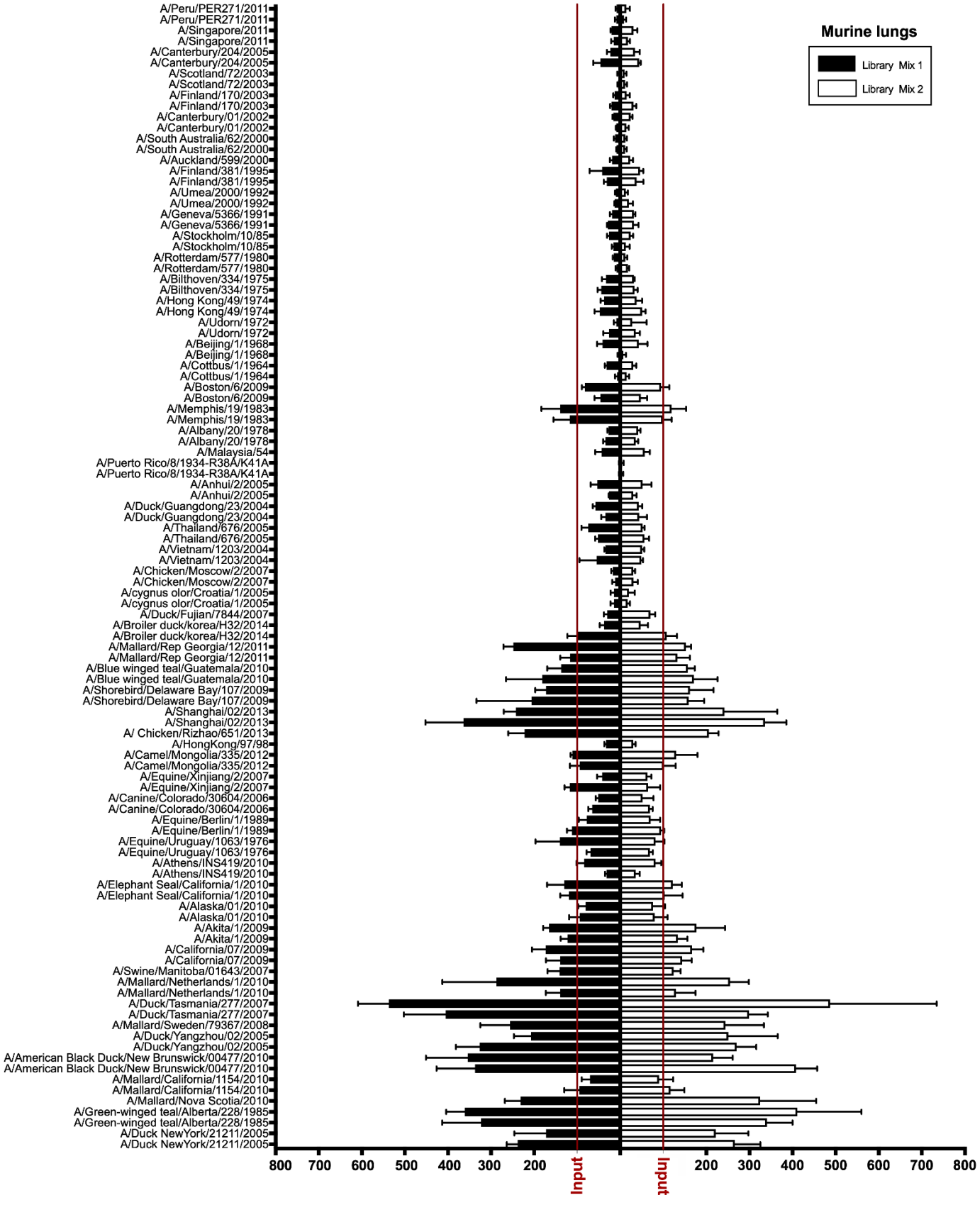

a
WT
Rag1-/-
Day0
Day1
Day3
Day4
Day5
Library Infection
Lung Tissue Collection
b
pfu/ml
Days post-infection
WT
Rag1-/-
c
Body weight loss (%)
Days post-infection
WT PBS
RAG1-/- PBS
WT Library
RAG1-/- Library
d
Percent survival
Days post-infection
WT PBS
Rag1-/- PBS
WT Library
Rag1-/- Library
e
Rag1-/-
WT
Input
Fold Induction vs Input (%)
A/Duck/New York/2005
A/GWT/Alberta/1985
A/Mallard/NS/2010
A/Mallard/California/2010
A/AmBlackDuck/NB/2010
A/Duck/Yangzhou/2005
A/Mallard/Sweden/2008
A/Duck/Tasmania/2007
A/Mallard/NL/2010
A/Swine/Manitoba/2007
A/California/07/2009
A/Akita/1/2009
A/Alaska/01/2010
A/Elephant seal/CA/2010
A/Athens/INS419/2010
A/Equine/Uruguay/1976
A/Equine/Berlin/1/1989
A/Canine/Colorado/2006
A/Equine/Xinjiang/2007
A/Camel/Mongolia/2012
A/Hong Kong/97198
A/Chicken/Rizhao/2013
A/Shanghai/02/2013
A/Shorebird/DE Bay/2009
A/BWT/Guatemala/2010
A/Mallard/R.Georgia/2011
A/Broilerduck/Korea/2014
A/Duck/Fujian/7844/2007
A/Cygnus/Croatia/2005
A/Chicken/Moscow/2007
A/Vietnam/1203/2004
A/Thailand/676/2005
A/Duck/Guangdong/2004
A/Annui/2/2005
A/PR8/1934 R38A/K41A
A/Malaysia/54
A/Albany/20/1978
A/Memphis/19/1983
A/Boston/6/2009
A/Cottbus/1/1964
A/Beijing/1/1968
A/Udorn/1972
A/Hong Kong/49/1974
A/Bilthoven/334/1975
A/Rotterdam/577/1980
A/Stockholm/10/1985
A/Geneva/5366/1991
A/Umea/2000/1992
A/Finland/381/1995
A/Auckland/599/2000
A/South Australia/62/2000
A/Canterbury/01/2002
A/Finland/170/2003
A/Scotland/72/2003
A/Canterbury/204/2005
A/Singapore/2011
A/Peru/PER271/2011
0.03

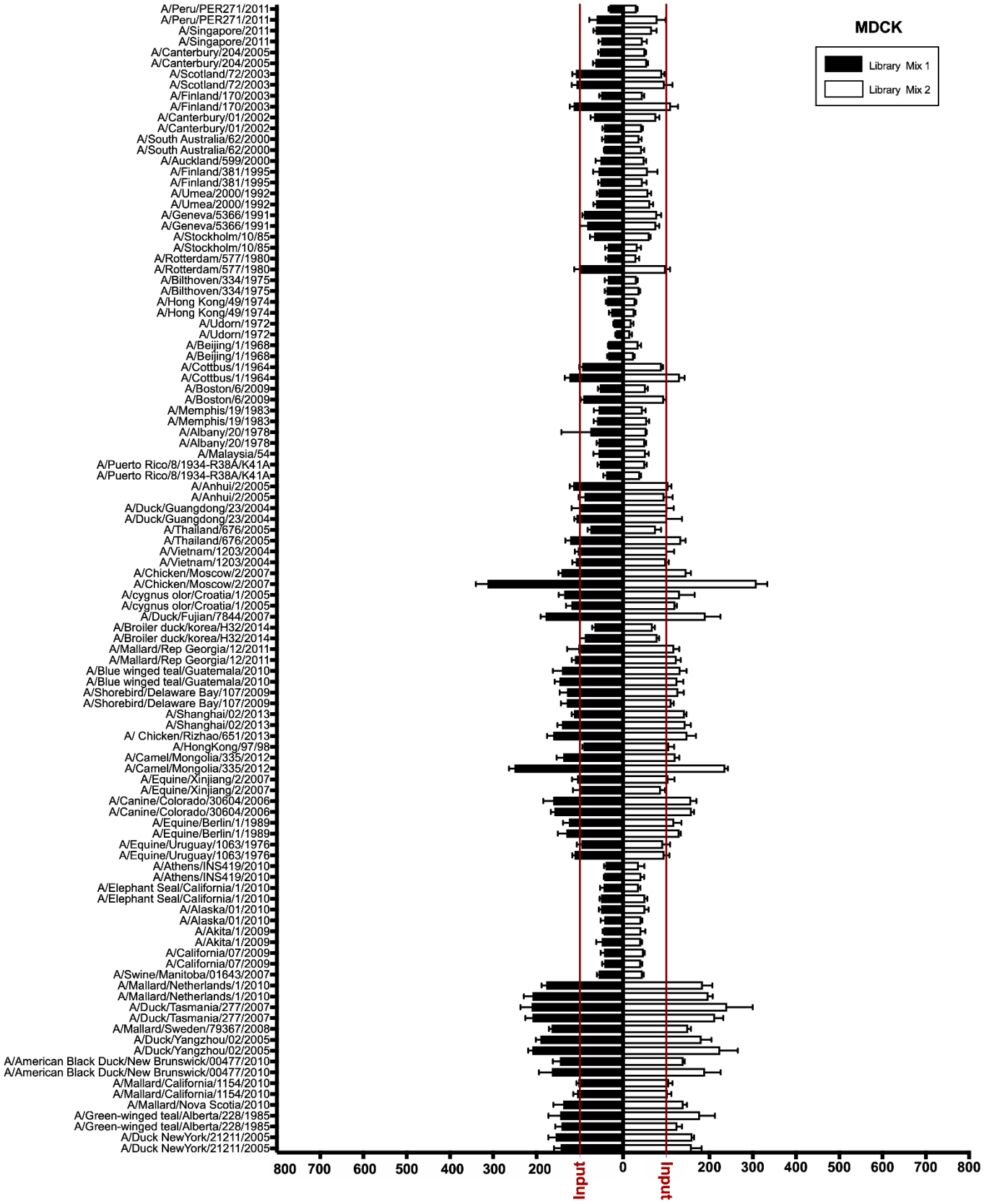
Figure S6, related to Figure 5. NS1 library selection profile is early host adaptive immune response independent. (**a**) Experimental layout of library infection in C57BL/6 wild type (WT) and C57BL/6 *Rag1*$^{-/-}$ mice. (**b**) Lung tissue from infected mice (n=3) were collected and homogenized at the indicated timepoints. Viral titers were determined by plaque assay on MDCK cells. Additionally, weight loss (**c**) and survival rates (**d**) were daily monitored during the infection. (**e**) Circular bar graph comparing results obtained from infecting WT or *Rag1*$^{-/-}$ at day 4 post infection and expressed as the relative barcode-fold increase percentage over the input. Error bars depict the standard deviation (SD).

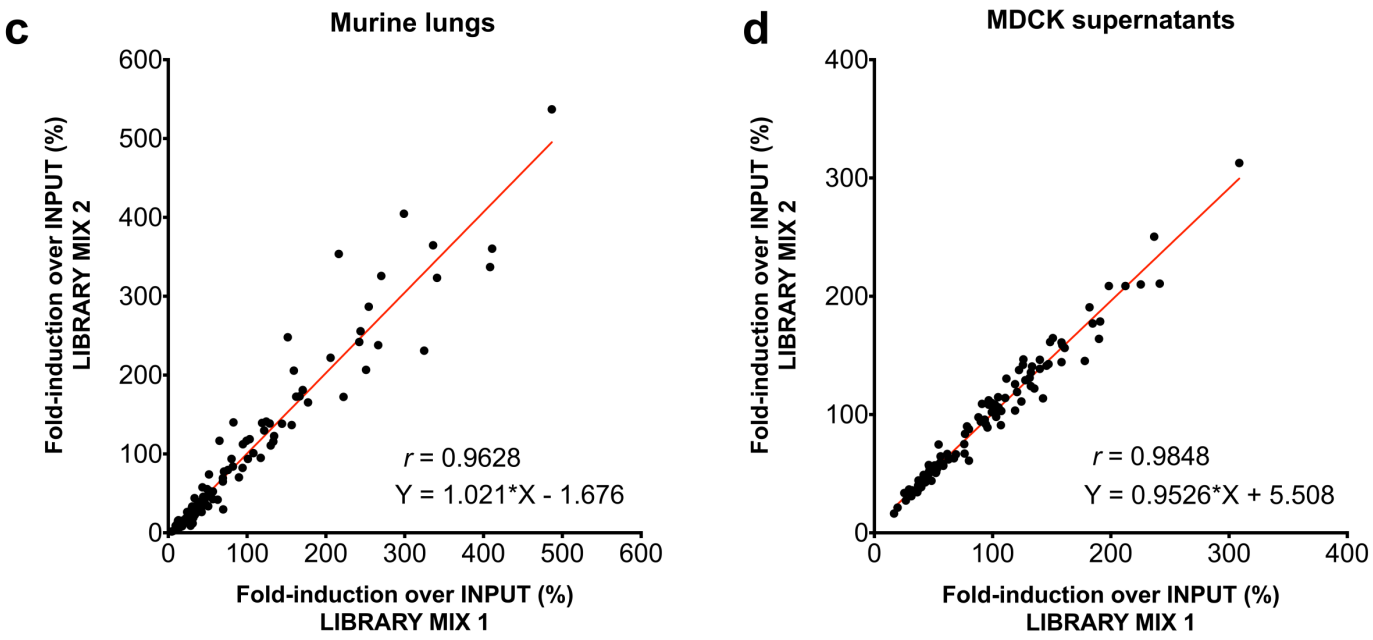
DATA S1

a


Murine lungs
Library Mix 1
Library Mix 2
A/Peru/PER271/2011
A/Peru/PER271/2011
A/Singapore/2011
A/Singapore/2011
A/Canterbury/204/2005
A/Canterbury/204/2005
A/Scotland/72/2003
A/Scotland/72/2003
A/Finland/170/2003
A/Finland/170/2003
A/Canterbury/01/2002
A/Canterbury/01/2002
A/South Australia/62/2000
A/South Australia/62/2000
A/Auckland/599/2000
A/Finland/381/1995
A/Finland/381/1995
A/Umea/2000/1992
A/Umea/2000/1992
A/Geneva/5366/1991
A/Geneva/5366/1991
A/Stockholm/10/85
A/Stockholm/10/85
A/Rotterdam/577/1980
A/Rotterdam/577/1980
A/Bilthoven/334/1975
A/Bilthoven/334/1975
A/Hong Kong/49/1974
A/Hong Kong/49/1974
A/Udorn/1972
A/Udorn/1972
A/Beijing/1/1968
A/Beijing/1/1968
A/Cottbus/1/1964
A/Cottbus/1/1964
A/Boston/6/2009
A/Boston/6/2009
A/Memphis/19/1983
A/Memphis/19/1983
A/Albany/20/1978
A/Albany/20/1978
A/Malaysia/54
A/Puerto Rico/8/1934-R38A/K41A
A/Puerto Rico/8/1934-R38A/K41A
A/Anhui/2/2005
A/Anhui/2/2005
A/Duck/Guangdong/23/2004
A/Duck/Guangdong/23/2004
A/Thailand/676/2005
A/Thailand/676/2005
A/Vietnam/1203/2004
A/Vietnam/1203/2004
A/Chicken/Moscow/2/2007
A/Chicken/Moscow/2/2007
A/cygnus olor/Croatia/1/2005
A/cygnus olor/Croatia/1/2005
A/Duck/Fujian/7844/2007
A/Broiler duck/korea/H32/2014
A/Broiler duck/korea/H32/2014
A/Mallard/Rep Georgia/12/2011
A/Mallard/Rep Georgia/12/2011
A/Blue winged teal/Guatemala/2010
A/Blue winged teal/Guatemala/2010
A/Shorebird/Delaware Bay/107/2009
A/Shorebird/Delaware Bay/107/2009
A/Shanghai/02/2013
A/Shanghai/02/2013
A/ Chicken/Rizhao/651/2013
A/HongKong/97/98
A/Camel/Mongolia/335/2012
A/Camel/Mongolia/335/2012
A/Equine/Xinjiang/2/2007
A/Equine/Xinjiang/2/2007
A/Canine/Colorado/30604/2006
A/Canine/Colorado/30604/2006
A/Equine/Berlin/1/1989
A/Equine/Berlin/1/1989
A/Equine/Uruguay/1063/1976
A/Equine/Uruguay/1063/1976
A/Athens/INS419/2010
A/Athens/INS419/2010
A/Elephant Seal/California/1/2010
A/Elephant Seal/California/1/2010
A/Alaska/01/2010
A/Alaska/01/2010
A/Akita/1/2009
A/Akita/1/2009
A/California/07/2009
A/California/07/2009
A/Swine/Manitoba/01643/2007
A/Mallard/Netherlands/1/2010
A/Mallard/Netherlands/1/2010
A/Duck/Tasmania/277/2007
A/Duck/Tasmania/277/2007
A/Mallard/Sweden/79367/2008
A/Duck/Yangzhou/02/2005
A/Duck/Yangzhou/02/2005
A/American Black Duck/New Brunswick/00477/2010
A/American Black Duck/New Brunswick/00477/2010
A/Mallard/California/1154/2010
A/Mallard/California/1154/2010
A/Mallard/Nova Scotia/2010
A/Green-winged teal/Alberta/228/1985
A/Green-winged teal/Alberta/228/1985
A/Duck NewYork/21211/2005
A/Duck NewYork/21211/2005
800 700 600 500 400 300 200
Input
Input
200 300 400 500 600 700 800

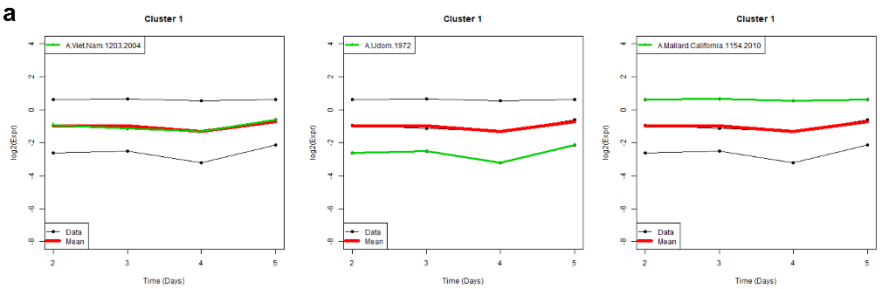
b

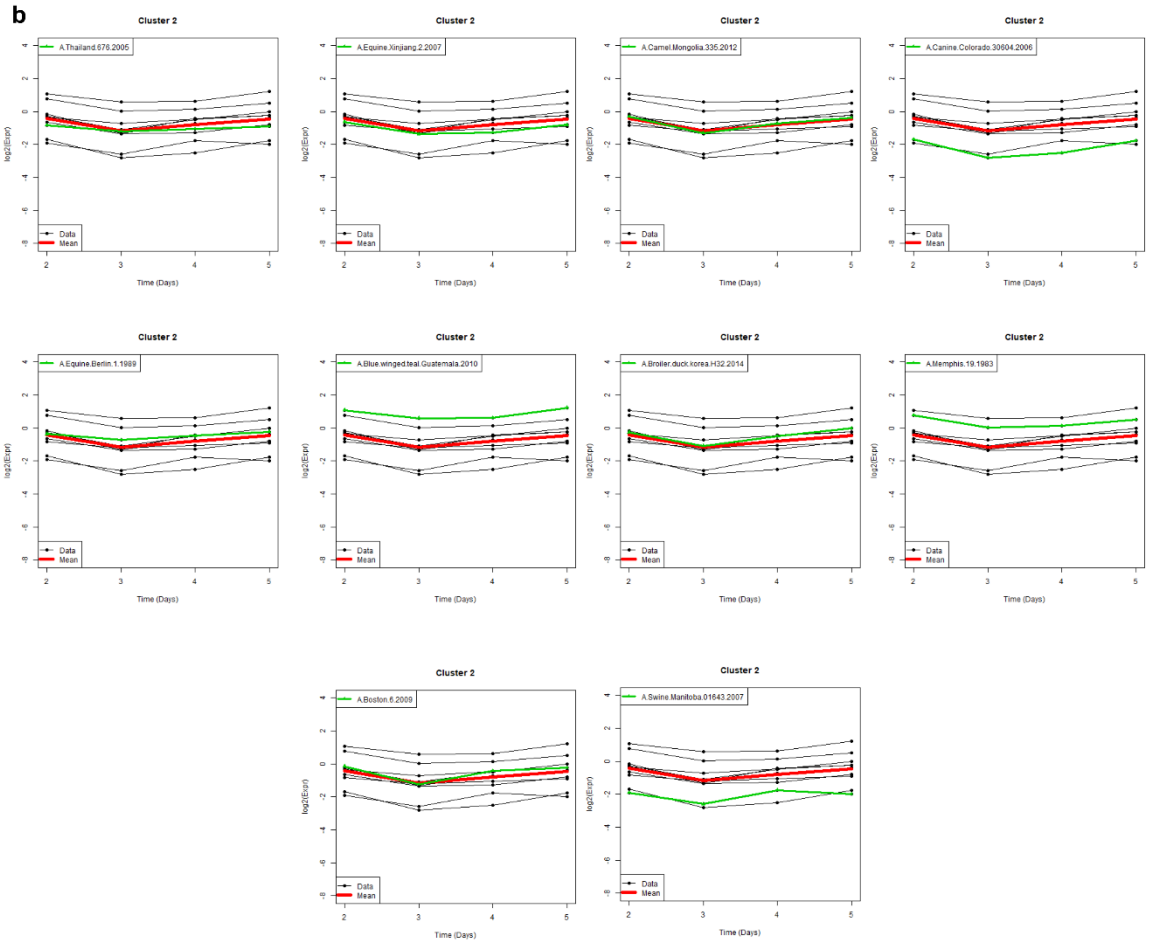

MDCK
Library Mix 1
Library Mix 2
A/Peru/PER271/2011
A/Peru/PER271/2011
A/Singapore/2011
A/Singapore/2011
A/Canterbury/204/2005
A/Canterbury/204/2005
A/Scotland/72/2003
A/Scotland/72/2003
A/Finland/170/2003
A/Finland/170/2003
A/Canterbury/01/2002
A/Canterbury/01/2002
A/South Australia/62/2000
A/South Australia/62/2000
A/Auckland/599/2000
A/Finland/381/1995
A/Finland/381/1995
A/Umea/2000/1992
A/Umea/2000/1992
A/Geneva/5366/1991
A/Geneva/5366/1991
A/Stockholm/10/85
A/Stockholm/10/85
A/Rotterdam/577/1980
A/Rotterdam/577/1980
A/Bilthoven/334/1975
A/Bilthoven/334/1975
A/Hong Kong/49/1974
A/Hong Kong/49/1974
A/Udorn/1972
A/Udorn/1972
A/Beijing/1/1968
A/Beijing/1/1968
A/Cottbus/1/1964
A/Cottbus/1/1964
A/Boston/6/2009
A/Boston/6/2009
A/Memphis/19/1983
A/Memphis/19/1983
A/Albany/20/1978
A/Albany/20/1978
A/Malaysia/54
A/Puerto Rico/8/1934-R38A/K41A
A/Puerto Rico/8/1934-R38A/K41A
A/Anhui/2/2005
A/Anhui/2/2005
A/Duck/Guangdong/23/2004
A/Duck/Guangdong/23/2004
A/Thailand/676/2005
A/Thailand/676/2005
A/Vietnam/1203/2004
A/Vietnam/1203/2004
A/Chicken/Moscow/2/2007
A/Chicken/Moscow/2/2007
A/cygnus olor/Croatia/1/2005
A/cygnus olor/Croatia/1/2005
A/Duck/Fujian/7844/2007
A/Broiler duck/korea/H32/2014
A/Broiler duck/korea/H32/2014
A/Mallard/Rep Georgia/12/2011
A/Mallard/Rep Georgia/12/2011
A/Blue winged teal/Guatemala/2010
A/Blue winged teal/Guatemala/2010
A/Shorebird/Delaware Bay/107/2009
A/Shorebird/Delaware Bay/107/2009
A/Shanghai/02/2013
A/Shanghai/02/2013
A/ Chicken/Rizhao/651/2013
A/HongKong/97/98
A/Camel/Mongolia/335/2012
A/Camel/Mongolia/335/2012
A/Equine/Xinjiang/2/2007
A/Equine/Xinjiang/2/2007
A/Canine/Colorado/30604/2006
A/Canine/Colorado/30604/2006
A/Equine/Berlin/1/1989
A/Equine/Berlin/1/1989
A/Equine/Uruguay/1063/1976
A/Equine/Uruguay/1063/1976
A/Athens/INS419/2010
A/Athens/INS419/2010
A/Elephant Seal/California/1/2010
A/Elephant Seal/California/1/2010
A/Alaska/01/2010
A/Alaska/01/2010
A/Akita/1/2009
A/Akita/1/2009
A/California/07/2009
A/California/07/2009
A/Swine/Manitoba/01643/2007
A/Mallard/Netherlands/1/2010
A/Mallard/Netherlands/1/2010
A/Duck/Tasmania/277/2007
A/Duck/Tasmania/277/2007
A/Mallard/Sweden/79367/2008
A/Duck/Yangzhou/02/2005
A/Duck/Yangzhou/02/2005
A/American Black Duck/New Brunswick/00477/2010
A/American Black Duck/New Brunswick/00477/2010
A/Mallard/California/1154/2010
A/Mallard/California/1154/2010
A/Mallard/Nova Scotia/2010
A/Green-winged teal/Alberta/228/1985
A/Green-winged teal/Alberta/228/1985
A/Duck NewYork/21211/2005
A/Duck NewYork/21211/2005
800 700 600 500 400 300 200 0 200 300 400 500 600 700 800
Input
Input

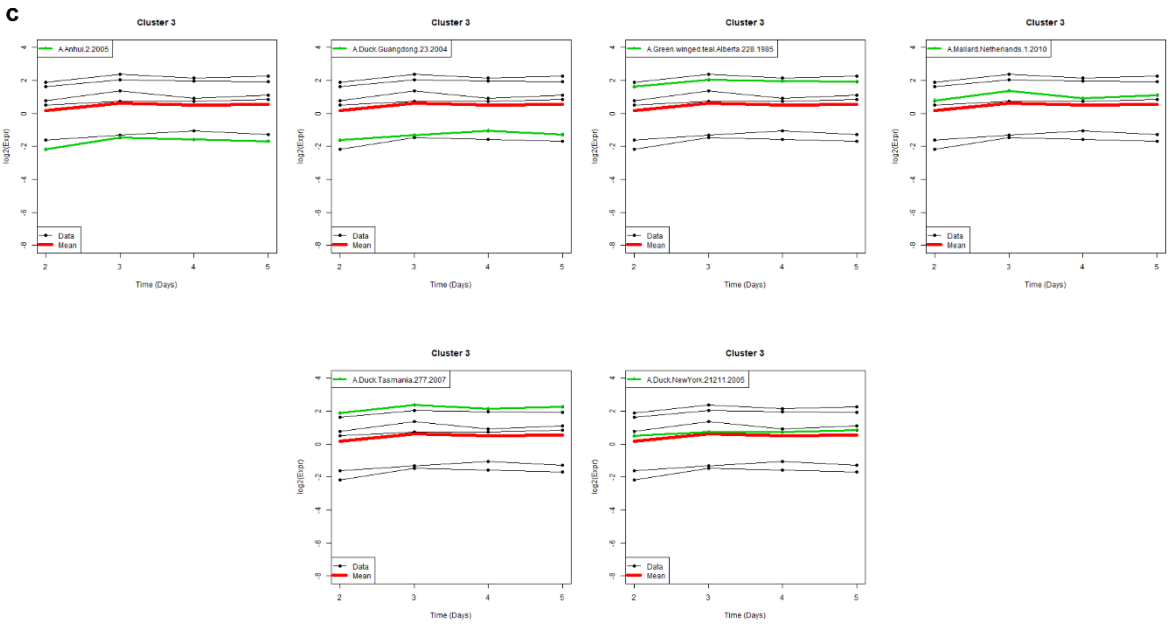


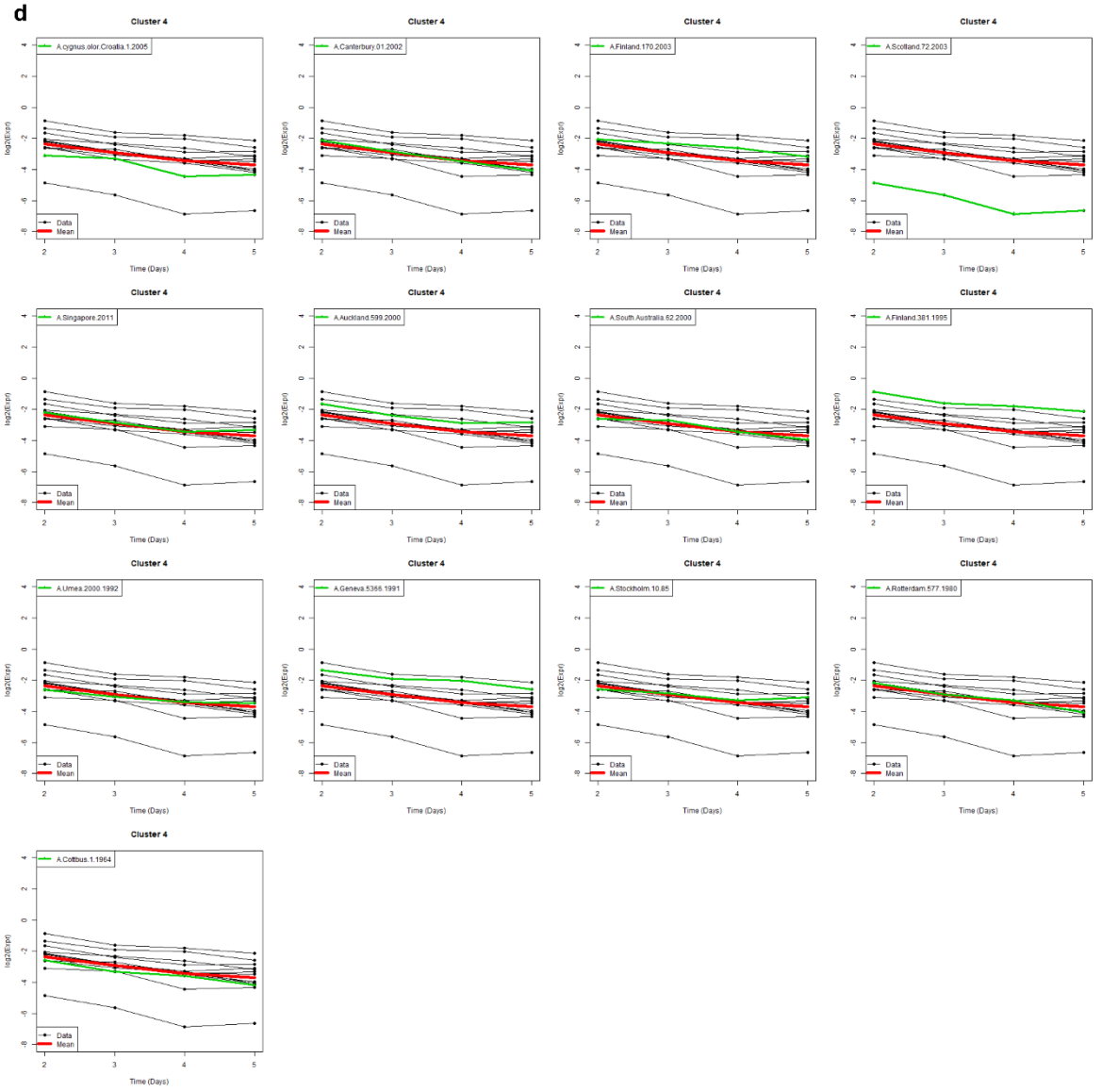
Data S1, related to Figure 2. Validating the reproducibility of our NS1 recombinant viral library *in vitro* and *in vivo*. Two independent viral library sets (100 pfu/virus) were used to infect 8-week-old female BALB/c mice (**a**) and MDCK cells (**b**) in triplicates. Relative barcode abundance was analyzed at 48 hours post-infection. Results are expressed as the fold-induction percentage over the initial relative proportion of barcode reads found in the initial viral mix (input, red dotted line). Scatter plots expressing the fold induction (%) over the input after infecting BALB/c mice (**c**) and MDCK cells (**d**) in triplicates with two different pools of library (Library Mix 1 or 2). Error bars depict the standard deviation (SD). Two-tailed paired t tests between independent library mixes were used to obtain Pearson correlation coefficients (r), with linear regression analysis.

DATA S2

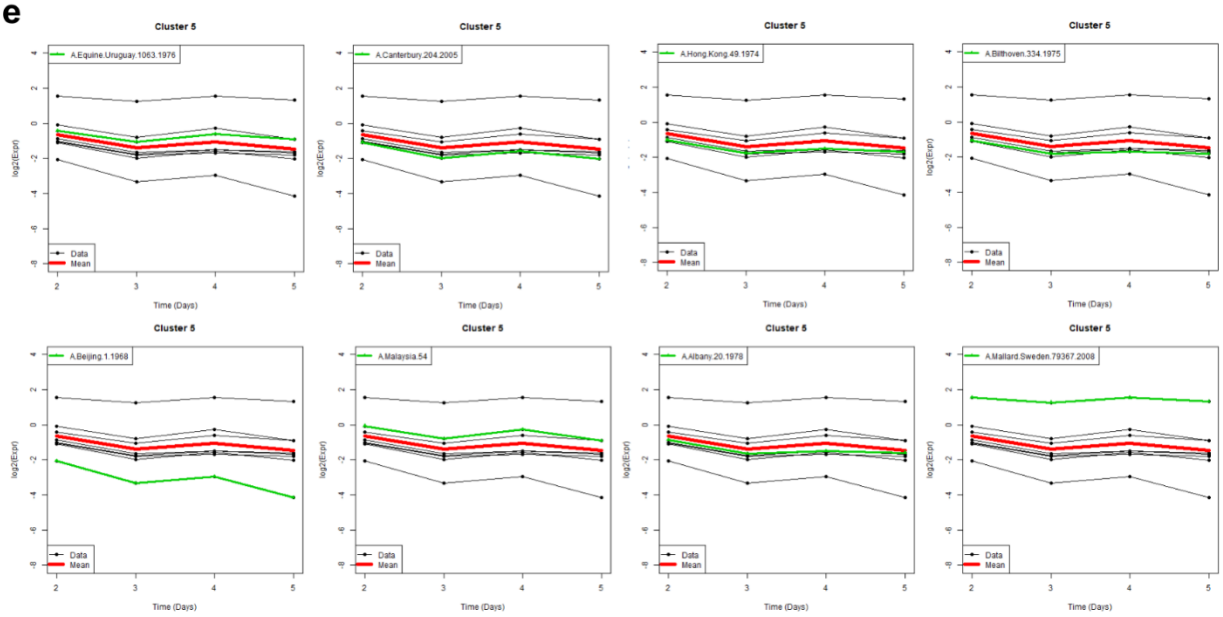

a
Cluster 1
A.Viet.Nam.1203.2004
A.Udorn.1972
A.Mallard.California.1154.2010
log2(Expr)
Time (Days)
Data
Mean


b
Cluster 2
A.Thailand.676.2005
A.Equine.Xinjiang.2.2007
A.Camel.Mongolia.335.2012
A.Canine.Colorado.30604.2006
A.Equine.Berlin.1.1989
A.Blue.winged.teal.Guatemala.2010
A.Broiler.duck.korea.H32.2014
A.Memphis.19.1983
A.Boston.6.2009
A.Swine.Manitoba.01643.2007
log2(Expr)
Time (Days)
Data
Mean

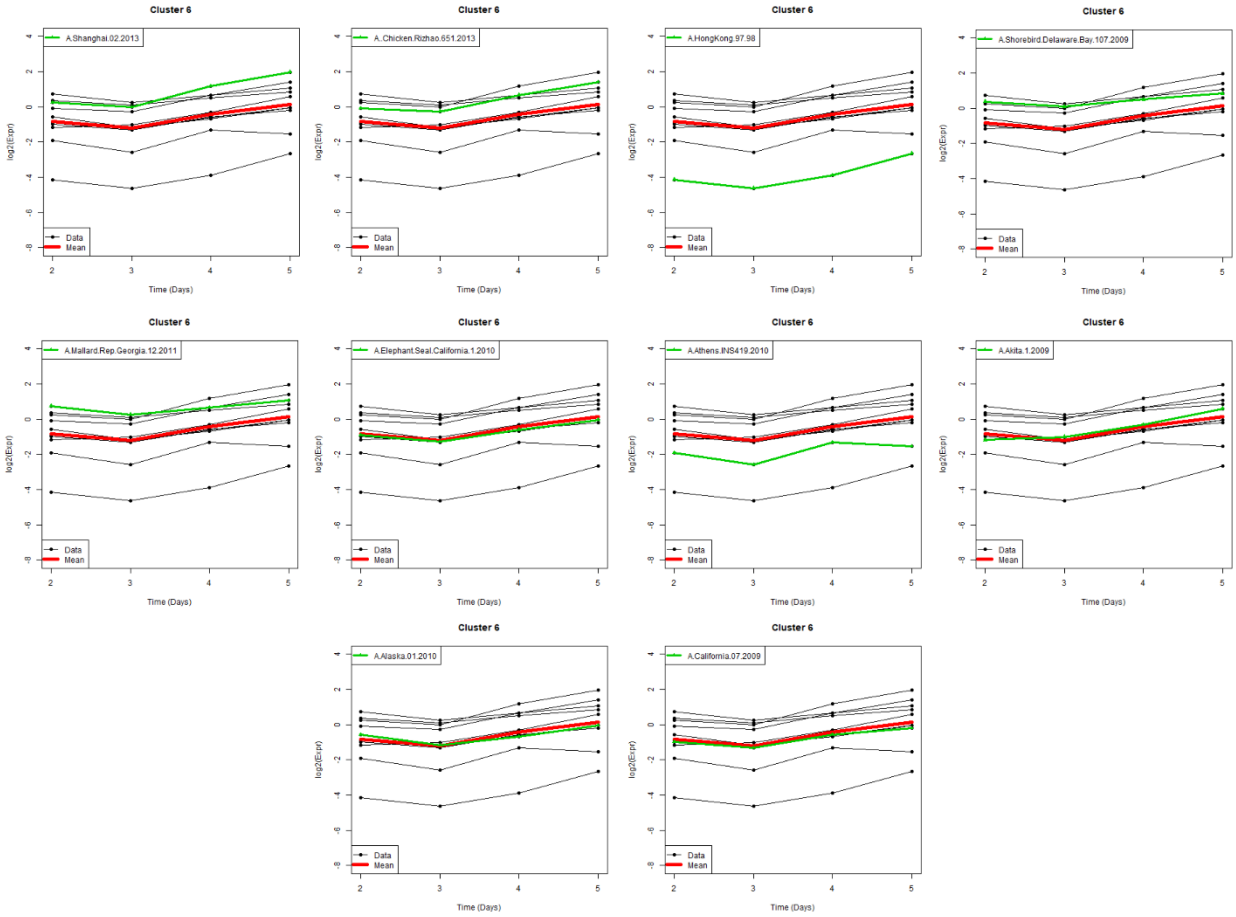

c
Cluster 3
A.Anhui.2.2005
A.Duck.Guangdong.23.2004
A.Green.winged.teal.Alberta.228.1985
A.Mallard.Netherlands.1.2010
A.Duck.Tasmania.277.2007
A.Duck.NewYork.21211.2005
Data
Mean
log2(Expr)
Time (Days)


d
Cluster 4
A.cygnus.olor.Croatia.1.2005
A.Canterbury.01.2002
A.Finland.170.2003
A.Scotland.72.2003
A.Singapore.2011
A.Auckland.599.2000
A.South.Australia.62.2000
A.Finland.381.1995
A.Umea.2000.1992
A.Geneva.5366.1991
A.Stockholm.10.85
A.Rotterdam.577.1980
A.Cottbus.1.1964
Data
Mean
log2(Expr)
Time (Days)

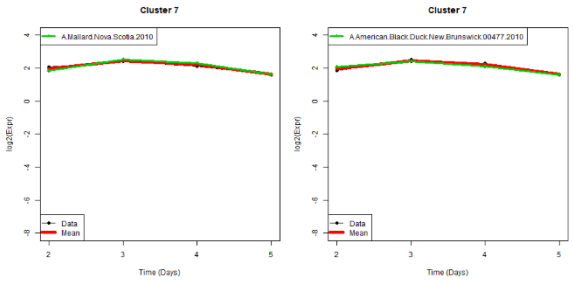
e


Cluster 5
A.Equine.Uruguay.1063.1976
A.Canterbury.204.2005
A.Hong.Kong.49.1974
A.Bilthoven.334.1975
A.Beijing.1.1968
A.Malaysia.54
A.Albany.20.1978
A.Mallard.Sweden.79367.2008
log2(Expr)
Time (Days)
Data
Mean

f


Cluster 6
A.Shanghai.02.2013
A..Chicken.Rizhao.651.2013
A.HongKong.97.98
A.Shorebird.Delaware.Bay.107.2009
A.Mallard.Rep.Georgia.12.2011
A.Elephant.Seal.California.1.2010
A.Athens.INS419.2010
A.Akita.1.2009
A.Alaska.01.2010
A.California.07.2009
log2(Expr)
Time (Days)
Data
Mean

g



Data S2, related to Figure 6. Time course analysis of clusters in Figure 6. Barcode reads for each recombinant virus within clusters 1 (**a**), 2 (**b**), 3 (**c**), 4 (**d**), 5 (**e**), 6 (**f**) and 7 (**g**) at days 2, 3, 4 and 5 post-infection were plotted (black arrows). Each panel highlights a data set of a particular virus (green) and the average values of the whole cluster (red).

TABLE S1

List of barcode sequences used in the viral library		
Barcode #	**Sequence**	**NS1 strain**
BC1	ATTTATATAAGCAGCTAACTGT	A/Udorn/1972
BC2	ATATTGATTAACAAACTAGCCA	A/Puerto Rico/8/1934
BC3	ATAGCTTACCGGACTCGGTGCC	A/Puerto Rico/8/1934 R38A/K41A
BC4	TAAAATTGCGATTCACTGGCCT	A/Shanghai/02/2013
BC5	TACAGTTAGCATAAGATGTGAG	A/Udorn/1972
BC6	TTTGTTTATATTTAACAGTGTG	A/Puerto Rico/8/1934
BC7	TTTCAGCATGAGAATCTCCTTC	A/Puerto Rico/8/1934 R38A/K41A
BC8	CTGCTCTTTGGCTGAATGGGCC	A/Shanghai/02/2013
BC9	AATTAATTTCATTGACTTCGGC	Akita/1/2009
BC10	ATAACCTGAAAGTACATATGCT	Akita/1/2009
BC11	ATAATGTCCCTCAATGTCGGCC	A/Alaska/01/2010
BC12	ATTTGGGCTGCCATGTCCAGGA	A/Alaska/01/2010
BC13	TATACAAATTATATTATGTCCG	A/Albany/20/1978
BC14	TTCTGTTAGAACTGCAGCATAT	A/Albany/20/1978
BC15	TTCAATCAATATCCAATTCCTC	A/Anhui/2/2005
BC16	CTTCAAAGTCTTCTAAGATCCG	A/Anhui/2/2005
BC17	AATAATCTCCACAGCCATCCAT	A/Athens/2010
BC18	TAAATCCATTATGTACTTGGTC	A/Athens/2010
BC19	TTTAACCCATTAATAATATGCA	A/Auckland/2000
BC21	TTTGGTGATGAATGACGCTGGC	A/Beijing/1/1968
BC22	ATTGATTTATTCAGTAACCGGA	A/Beijing/1/1968
BC23	AATATCAACAAGTAGTCCCGCT	A/Bilthoven/334/1975
BC24	TTAAACTTTCTGGGAGATGGCC	A/Bilthoven/334/1975
BC25	TTAAGCCAGCTAATAAAGAGTC	A/Blue winged teal/Guatemala/2010
BC26	TTATATAAACTCAGGACCGCCA	A/Blue winged teal/Guatemala/2010
BC27	TTACACCACTACATTGAAGGAT	A/Boston/6/2009
BC28	TTTGGTTTGCCTGGAGATGCCA	A/Boston/6/2009
BC31	TATATCTGTGTTTGTCGGAGAC	A/Broiler_duck/korea/2014
BC32	TTAAATTCTCCACTCAGGTGGA	A/Broiler_duck/korea/2014
BC33	ACAATTCCCTTCTGCAAGCGGT	A/California/07/2009
BC34	TTCATCAAAGATATTCTTGGTG	A/California/07/2009
BC35	ATTTCAAATAACAGATTAAGTT	A/Camel/Mongolia/2012
BC36	TTCATATGGCCCGTGCTTAGCC	A/Camel/Mongolia/2012
BC37	TATAGTGATGCCGTTGTCAGTG	A/Canine/Colorado/2006
BC38	ATAACCCTTATGTTCAAATGGA	A/Canine/Colorado/2006
BC39	TTAACTTCTCTTCTTGTAAGGC	A/Canterbury/01/2002
BC40	TTCTGTAATAATATTCCAGGCC	A/Canterbury/01/2002
BC41	AATGTTGAACACAAGGTTGCTC	A/Canterbury/204/2005
BC42	TTCTTAAATAAGAATAAGAGCC	A/Canterbury/204/2005
BC43	GTCTAATACGATAATAAGCCGG	A/Moscow/2007
BC44	TTTCTCGTGATACTTCAATGGG	A/Moscow/2007
BC47	TTTGGCGATGATAATTGTAGGG	A/ Rizhao/2013
BC48	ATAAACAAGAATTAGATATCTT	A/ Rizhao/2013
BC49	TTAATGGAGAGCTGGCTGGCCT	A/Cottbus/1964
BC50	TAGCTCATAGATGTAGTGTCGT	A/Cottbus/1964
BC51	TATATTTCAGCATGACAGACCA	A/Cygnus olor/2005
BC52	TTTCTGGATGGATCACTGGGTG	A/Cygnus olor/2005

BC53	ATTCATTAATCTTTACAGTGCG	A/Fujian/2007
BC54	TATTGTGTTAATTATGGATGAA	A/Fujian/2007
BC55	TATACCTCTGAGGTTTCTTCCA	A/Guangdong/2004
BC56	TTAACGTAGACTTTCAGCTGCA	A/Guangdong/2004
BC57	ATCATCTGACTCATGATAGGTC	A/Equine/Berlin/1989
BC58	TTACTTCATTCACTTGTGTGTT	A/Equine/Berlin/1989
BC59	ATTCGTCATAATTCATGATCTT	A/Equine/Uruguay/1976
BC60	TTATTATCTTGGATAAGGAGGC	A/Equine/Uruguay/1976
BC61	TAACAGTTACTTATCTAATCCC	A/Equine/Xinjiang/2007
BC62	TTTCTGGTAATGACGAAGGGTC	A/Equine/Xinjiang/2007
BC63	TACAATCTCGATCTTACTGCGA	A/Finland/2003
BC64	CATTCATCTCCATTGCATTGGA	A/Finland/2003
BC65	TAGTGAAGCCACAGATGTA	A/Finland/95
BC66	ATATTACATAAATTTATCATGC	A/Finland/95
BC69	TAATTGTGGATGGATTGGAGAT	A/Geneva/1991
BC70	TTTCCGATCATCCTGAAGAGGC	A/Geneva/1991
BC71	TTAATTATCTTTCAAACGAGGA	A/Hong Kong/1974
BC72	ATAATCTTCCTGCGAATGTGGG	A/Hong Kong/1974
BC75	TTGCTTTAAATCTCTCCTTGGA	A/Malaysia/54
BC77	TTTCTAAGTAACCAACATAGCC	A/Rep Georgia/2011
BC78	ATCTGATATGGCAATCTTTCCT	A/Rep Georgia/2011
BC79	TAAGTCCAAACTCCTTAACAGC	A/Memphis/1983
BC80	TTAATATGTCAAATCCATATTG	A/Memphis/1983
BC81	ATAACTCCTTACATTTACCCAA	A/Peru/2011
BC82	ATATGTAAACTCTAGGACGTTG	A/Peru/2011
BC83	ATAAGGTGACACAGGAACATAC	A/Rotterdam/1980
BC84	ATATCTGGAGTAAGAAACAGAG	A/Rotterdam/1980
BC85	ATTAAAAGCATCATAATCAGGA	A/Scotland/2003
BC86	AAGATGAGACATCTTGTAAGCA	A/Scotland/2003
BC89	TTTCCAGTCACTTTTGTCTGCT	A/Shorebird/Delaware/2009
BC90	TTAGAGTAATCTCCAATCAGCT	A/Shorebird/Delaware/2009
BC91	TTGAAATTTCCTGTAGGAGATT	A/Singapore/2011
BC92	TTAGCGCTGGAGCCCAGGTGAC	A/Singapore/2011
BC93	TTACTTAAAGCCGGTGGTGCCG	A/South Australia/2000
BC94	TTCCAGTTTAGTTATCTGAGCG	A/South Australia/2000
BC95	TAATAGCTAAGTTATTACAGGC	A/Stockholm/85
BC96	TTTGAAAGCAAATTAAATACAG	A/Stockholm/85
BC97	TTGCCATGGATAAGTAATTGGT	A/Manitoba/2007
BC98	TAATGCTAACAACAACAAACCA	A/Manitoba/2007
BC101	ATGGAAATTACAAGTAGGAGCC	A/Thailand/2005
BC102	TTAGTTCCTGCCTGCTTGAGGT	A/Thailand/2005
BC103	TTAATAATTCATTTAGGAGCCG	A/Umea/92
BC104	AATTGTATACATCCTCCGAGGG	A/Umea/92
BC105	ATTACTTTGAGAACTTAAAGCA	A/Black Duck/New Brunswick/2010
BC106	ATAAGTTTCGTCATAACTCGCC	A/Black Duck/New Brunswick/2010
BC107	TTCCAGGTTACATACGACTGCG	A/Duck New York/21211/2005
BC108	TAATGCCAGGTGAAATTCTCTC	A/Duck New York/21211/2005
BC109	TTGTGCCTACTGCCTCGGAATT	A/Tasmania/2007
BC110	TTATCTTCTCAGCCAGATCCGT	A/Tasmania/2007
BC111	TTCTCTTAAAGGGCAGTGTCAG	A/Duck/Yangzhou/02/2005
BC112	TATGAGAATAAGGATGAGAGGT	A/Duck/Yangzhou/02/2005
BC113	TACAATGATAAGAGAAGCACCA	A/Mallard/California/1154/2010
BC114	ATATTACAAGGCTTGGAAAGGT	A/Mallard/California/1154/2010

BC115	AAGTCGCCCTACGGCGGGTGCC	A/green-winged teal/1985
BC116	TTTCATATATGGGTTCTCACAG	A/green-winged teal/1985
BC117	AATAGTCCAGAATTTCACTGGC	A/Mallard/Netherlands/2010
BC118	TTAATAGCATCGTATTTGTCCT	A/Mallard/Netherlands/2010
BC120	TTACCACAAGAAATAAAGACCA	A/Mallard/Nova Scotia/2010
BC121	TTAAGATAAACAGGATTCAGCC	A/Mallard/Sweden/2008
BC123	ATAGAGGTCAGAGAGTCGTCCC	A/Elephant Seal/California/1/2010
BC124	GACATTCATAACAGCAAATGGC	A/Elephant Seal/California/1/2010
BC125	TATTAGTCATACATCTTCCTTG	A/Hong Kong/97/98
BC127	TTTCTCAAACAGCTGGTAACGC	A/Vietnam/1203/2004
BC128	AATATAAGTATCTTTAGCGGCG	A/Vietnam/1203/2004

Table S1, related to STAR Methods. List of barcode sequences used in the viral library.

TABLE S2

List of primers used for the deep-sequencing analysis	
Primer Name	**Sequence**
3-Ambi-NS	GATCGCTCTTCTGGGAGCaAAAGCAGGgtgac
5-SAP-NS	CATCGCTCTTCTATTAGTAGAAACAAGGgtgtt
Ilummina_mir30_F	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTAGTGAAGCCACAGATGTA
Ilummina_mir30_R_1	CAAGCAGAAGACGGCATACGAGATaaaatcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_2	CAAGCAGAAGACGGCATACGAGATaccgccGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_3	CAAGCAGAAGACGGCATACGAGATacgttaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_4	CAAGCAGAAGACGGCATACGAGATagagagGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_5	CAAGCAGAAGACGGCATACGAGATagtagaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_6	CAAGCAGAAGACGGCATACGAGATatgggtGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_7	CAAGCAGAAGACGGCATACGAGATcaatggGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_8	CAAGCAGAAGACGGCATACGAGATcataatGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_9	CAAGCAGAAGACGGCATACGAGATcgctaaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_10	CAAGCAGAAGACGGCATACGAGATcggcgcGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_11	CAAGCAGAAGACGGCATACGAGATctgacgGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_12	CAAGCAGAAGACGGCATACGAGATgattcaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_13	CAAGCAGAAGACGGCATACGAGATgcctgtGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_14	CAAGCAGAAGACGGCATACGAGATgcgcagGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_15	CAAGCAGAAGACGGCATACGAGATgtacgaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_16	CAAGCAGAAGACGGCATACGAGATgtcaacGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_17	CAAGCAGAAGACGGCATACGAGATtaccttGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_18	CAAGCAGAAGACGGCATACGAGATtaggacGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_19	CAAGCAGAAGACGGCATACGAGATtcaacaGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA
Ilummina_mir30_R_20	CAAGCAGAAGACGGCATACGAGATtctgtgGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCCAAAGTGATTTAATTTATACCATTTTA

Table S2, related to STAR Methods. List of primers used for the deep-sequencing analysis.