NETWORK-GUIDED DISCOVERY OF INFLUENZA VIRUS REPLICATION HOST FACTORS

Emily E. Ackerman[1], Eiryo Kawakami[2,3], Manami Katoh[3,4], Tokiko Watanabe[2,3], Shinji Watanabe[2], Yuriko Tomita[2,3], Tiago J. Lopes[2,5], Yukiko Matsuoka[3,4], Hiroaki Kitano[2,3,6,7], Jason E. Shoemaker[1-3,8,*,†], Yoshihiro Kawaoka[2,3,5,9,†]

[†]Shared senior authorship.

1 Department of Chemical & Petroleum Engineering, Swanson School of Engineering, University of Pittsburgh, Pittsburgh, PA, 15213, USA

2 Division of Virology, Department of Microbiology and Immunology, Institute of Medical Science, University of Tokyo, Tokyo 108-8639, Japan

3 ERATO Infection-Induced Host Responses Project, Japan Science and Technology Agency, Saitama 332-0012, Japan

4 The Systems Biology Institute, Minato-ku, Tokyo 108-0071, Japan

5 Department of Pathobiological Sciences, School of Veterinary Medicine, University of Wisconsin-Madison, 575 Science Drive, Madison, WI 53711, USA

6 Laboratory for Disease Systems Modeling, RIKEN Center for Integrative Medical Sciences, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan

7 Okinawa Institute of Science and Technology, Onna-son, Okinawa 904-0495, Japan

8 Department of Computational and Systems Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, 15213, USA

21

22 9 Department of Special Pathogens, International Research Center for Infectious Diseases,

23 Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108-8639, Japan

24 *Corresponding author: jason.shoemaker@pitt.edu

25

26

ABSTRACT

The position of host factors required for viral replication within a human protein-protein interaction (PPI) network can be exploited to identify drug targets that are robust to drug-mediated selective pressure. Host factors can physically interact with viral proteins, be a component of pathways regulated by viruses (where proteins themselves do not interact with viral proteins) or be required for viral replication but unregulated by viruses. Here, we demonstrate a method of combining a human PPI network with virus-host protein interaction data to improve antiviral drug discovery for influenza viruses by identifying target host proteins. Network analysis shows that influenza virus proteins physically interact with host proteins in network positions significant for information flow. We have isolated a subnetwork of the human PPI network which connects virus-interacting host proteins to host factors that are important for influenza virus replication without physically interacting with viral proteins. The subnetwork is enriched for signaling and immune processes. Selecting proteins based on network topology within the subnetwork, we performed an siRNA screen to determine if the subnetwork was enriched for virus replication host factors and if network position within the subnetwork offers an advantage in prioritization of drug targets to control influenza virus replication. We found that the subnetwork is highly enriched for target host proteins – more so than the set of host factors that physically interact with viral proteins. Our findings demonstrate that network positions are a powerful predictor to guide antiviral drug candidate prioritization.

IMPORTANCE

Integrating virus-host interactions with host protein-protein interactions, we have created a method using these established network practices to identify host factors (i.e. proteins) that

are likely candidates for antiviral drug targeting. We demonstrate that interaction cascades between host proteins that directly interact with viral proteins and host factors that are important to influenza replication are enriched for signaling and immune processes. Additionally, we show that host proteins that interact with viral proteins are in network locations of power. Finally, we demonstrate a new network methodology to predict novel host factors and validate predictions with an siRNA screen. Our results show that integrating virus-host proteins interactions is useful in the identification of antiviral drug target candidates.

INTRODUCTION

Viruses such as influenza virus hijack and reprogram host cellular machinery to perform virus replication tasks. Influenza outbreaks have a major impact on public health and the global economy each year(1, 2). While annual vaccinations provide some protection, vaccination effectiveness is impaired by antigenic drift and availability issues(3, 4). Recent sporadic human infections with avian viruses of H5N1 and H7N9 subtypes have raised concerns about the pandemic potential of these viruses(5–8). Antiviral drugs that target influenza viral proteins are available(9, 10) but drug resistant strains have emerged(11, 12). Therefore, there is an urgent need to identify drug targets that are robust to virus mutation and drug-mediated selective pressure.

Understanding virus-host interactions in the context of the human protein-protein interaction (PPI) network will provide a global perspective into how influenza virus manipulates host proteins and aid in identifying host factors involved in influenza virus replication for drug targeting(13–15). The virus-host interactome is visualized in Fig. 1A. Within a PPI network, a protein's global significance can be assessed by the protein's

network centrality: the identification of important components based on information flow across the network. Common measures include a proteins degree (number of binding partners) and betweenness (the degree to which the protein is a bottleneck in the network) though several others exist(16, 17). Network centrality has emerged as a valuable tool for studying proteins associated with cancer(18, 19) and drug targeting(19–22). PPI network-based approaches have recently been utilized in influenza virus studies to identify and study potential factors involved in virus replication(23–27). Network studies have demonstrated that virus interacting host proteins tend to have a high network significance based on a variety of network metrics (including betweenness and degree) for several viruses including influenza viruses(28) and hepatitis C virus(29). A comparative analysis of influenza virus protein and host protein interactomes has identified novel host factors that are common across the protein interactomes(30). Furthermore, meta-analysis studies have developed statistical methods to better identify host factors by leveraging data from several virus replication screens (31). Yet, a remaining question is how effectively can virus-host protein interaction data and network topology be leveraged to improve host factor identification (i.e. antiviral drug target identification).

Here, we demonstrate a method of integrating virus-host protein interaction data into a human PPI network to prioritize host proteins as antiviral drug target candidates. First, we analyzed a set of 1,292 human proteins identified previously as having interactions with influenza virus proteins(32), 299 of which were found to significantly inhibit influenza virus replication during an siRNA virus replication screen (Fig. 1A). Consistent with previous studies, we show that virus-interacting human proteins tend to be in positions essential to PPI network information flow and are closely clustered within the PPI network. We then isolated the subnetwork of the human PPI network that connects virus-interacting host

proteins to non-interacting, host factors (referred to as "internal") that were identified to be important for influenza virus replication in a study and re-evaluated in this work (33) (Fig. 1B). Candidate proteins connecting virus-interacting host proteins to internal host factors were selected based on their betweenness within this subnetwork and evaluated by viral replication screen. Betweenness was selected under the hypothesis that selecting network bottlenecks (i.e. high betweenness proteins) would limit the opportunity for the virus to engage host machinery through alternative pathways. The fraction of proteins tested which significantly reduced virus replication (i.e. the hit rate) was compared to the hit rate observed in a genome-wide screen, the hit rate when screening virus-interacting proteins (the virus' nearest neighbors in the network) and the hit rate observed when screening host factor identified in a previous study(33).

RESULTS

Host proteins that interact with influenza virus proteins are central to the PPI network

Studies have shown that proteins in network positions that are essential for information flow within a PPI network (e.g. high degree or high betweenness) are more likely to be associated with diseases(34, 35) or drugs with known, dangerous side-effects(19, 36). Using a human PPI network, we analyzed the network topology characteristics of virus-interacting and non-virus-interacting host proteins. In a previous study, we identified 1,292 host proteins that co-precipitated with at least one of 11 influenza virus proteins (viral PB2, PB1, PA, HA, NP, NA, M1, M2, NS1, NS2, and PB1-F2 proteins)(32). These proteins are referred to as "virus-interacting proteins". We mapped the interaction data onto a human PPI network developed from the Human Integrated Protein-Protein Interaction rEference (HIPPIE) database(37). After constraining the network to highly confident interactions (see

Methods), the PPI consisted of one large network (9,969 proteins and 57,615 interactions) which contained 1,213 influenza virus-interacting host proteins and 86 smaller networks that contained 7 or fewer proteins (the majority only containing 2 proteins) and no influenza virus-interacting proteins. The smaller networks were removed from further consideration.

Virus proteins were significantly more likely to interact with host proteins that were in positions of high regulatory importance in the human PPI network. For every protein, the degree (number of neighbor proteins) and betweenness(38) (measure of the shortest paths between all other proteins in the network that include the protein in question) were calculated. On average, the degree of virus-interacting host proteins was twice the degree of all proteins in the network (Fig. 2A; the median degree of virus-interacting proteins = 10; the median degree of all proteins in the network = 5; Student t-test of the log-scaled data $p < 10^{-16}$). Virus-interacting proteins also had a significantly higher betweenness (Fig. 2B; virus-interacting proteins median betweenness = 1625.1; the median betweenness of all protein in the network = 32.8; Mann-Whitney U test of the log-scaled data $p < 10^{-16}$). We also compared to the median betweenness when removing proteins with a betweenness of zero. Virus-interacting proteins still had a significantly higher betweenness but the population medians were closer in value (virus-interacting proteins median betweenness = 3981.1; the median betweenness of all protein in the network = 1584.8; Mann-Whitney U test of the log-scaled data $p = 8.2*10^{-16}$). The tendency for virus proteins to bind host proteins that had a higher degree and betweenness was consistent when analyzing the interaction partners of each virus protein separately (Fig. S1; pairwise t-test of the log-scaled data. All $p < 0.01$ except for betweenness of NS2-interacting proteins which was not significantly distinct from the betweenness of the full PPI). This indicates that influenza virus proteins selectively

interact with human proteins that can strongly regulate cellular behavior. These results are consistent with literature findings for HCV and Dengue virus(39, 40) and with a previous study which used a yeast two-hybrid approach to identify influenza virus interacting host proteins for 10 of the 11 virus proteins (28). Further, these are characteristics that generalize to each virus protein's interacting partner; suggesting that all 11 virus proteins have a role in manipulating cellular machinery.

Influenza virus-interacting host proteins are closely connected in the human PPI network

Next, we evaluated if virus-interacting proteins tend to cluster closely to one another in the PPI network. A previous study suggested that host factors of viral replication are closely clustered within the network but did not assess the topological characteristics of virus-interacting host proteins (41). Functionally related proteins are often observed to be closely clustered in PPI networks(42, 43). Knowing that influenza virus proteins manipulate multiple host cell functions to promote replication, these previous studies suggest that the interaction partners of viral proteins should be closely clustered by host function. If true, neighboring cluster proteins could serve as possible alternatives for influenza virus to manipulate each host function.

We quantified how close each virus proteins' interacting host proteins are within the network by calculating the shortest distances required to connect all of the host proteins that interact with a viral protein, creating a distribution of distances. The cumulative distribution details the fraction of host proteins that could be connected to other host proteins that bind the same viral protein in *n* or fewer steps. As a control, we determined the cumulative distribution of distances that result from randomly sampled proteins in the network. For a single iteration, we created a set of random proteins. The size of the set was

167 determined by the number of proteins which interact with the virus protein of interest (e.g.
168 PB1 has 322 interacting host proteins, therefore 322 proteins were randomly selected from
169 the network; Fig. 2C-F). The distributions of distances connecting all of the randomly
170 sampled proteins was calculated. This was process was repeated 100 times.

171 We found that virus-interacting host proteins are very significantly clustered within the PPI
172 network. The set of proteins that interact with a viral protein are significantly more closely
173 clustered in the network than expected by chance (Fig. 2C-F, $p < 0.01$ when comparing the
174 median distance of the virus-interacting proteins to the median distance of randomly
175 sampled proteins). Generally, ~25% of the randomly sampled proteins are connected by 2
176 or fewer interactions while 88.7% of PB1-interacting proteins, 90.0% of HA-interacting
177 proteins, 98.2% of NS1-interacting proteins, and 79.6% of all host proteins that interact any
178 influenza virus protein are connected by 2 or fewer interactions. Collectively, these results
179 support that viral proteins are selectively targeting closely clustered host proteins.

180 We next evaluated if influenza interacting proteins are often components of a common
181 protein complex. To do so, we determined the fraction of all influenza virus interacting
182 proteins pairs (735,078 pairs in total) that appear within a protein complex and compared
183 that fraction to the fraction of all protein pairs (49,685,496 total pairs) in the PPI that appear
184 in a protein complex. Mammalian protein complex information was downloaded from
185 CORUM (a comprehensive resource of mammalian protein complex data)(44). We found
186 that 1.5% of all virus interacting protein pairs are involved in a complex where as only
187 0.066% of all proteins pairs in the PPI are involved in a complex. In sum, influenza virus
188 proteins are closely clustered and 22.4 times more likely to be involved in a protein complex
189 than randomly selected proteins.

Constructing the influenza virus-host subnetwork

Network analysis of virus-interacting host proteins demonstrates that viral proteins preferentially interact with closely connected host proteins that are in positions central to information flow across the human PPI network; suggesting that it may be possible to exploit network positions to prioritize potential antiviral drug targets. We hypothesized that there exists a subnetwork consisting of pathways that *connect* virus interacting proteins to key cellular machinery that is likely to be significantly enriched for host factors. We further hypothesized that the topology of host factors within this subnetwork may provide an additional advantage in identifying host factors.

To evaluate these hypotheses, we first performed an siRNA screen of host factors identified in a previous genome-wide screen for influenza virus host factors to identify key host factors that do not interact directly with the virus (33). Poor repeatability due to differences in the experimental conditions and possibly high false negative rates (41) often characterizes siRNA screens of influenza virus replication host factors. Here, HEK293 cells were transfected with siRNAs targeting 264 non-virus interacting host factors identified in Karlas et al 2010 (two siRNAs per gene were used, as shown in Table S1; AllStars Negative Control siRNA [QIAGEN] was used as a negative control), then infected with influenza virus at 24 hours post-transfection. The culture supernatants were harvested for virus titration at 48 hours post-infection. Virus titers were determined by plaque assay. A protein was defined as a hit if the virus titers decreased by at least two log units upon transfection with an adjusted $p < 0.01$. Cell viability of siRNA-transfected cells was assessed using Cell-Titer Glo assay and down-regulation of mRNA levels for the hit proteins were confirmed by qRT-PCR. Of the 264 previously identified host factors tested, 71 significantly down-regulated virus replication. Of the 71, 21 were identified to directly interact with influenza virus proteins. In all, 50 of

the host factors down-regulated virus growth and do not directly interact with the virus. We labeled these proteins as "internal-essential" host factors.

Next, we constructed an influenza virus specific subnetwork (process illustrated in Fig. 1B) of the shortest paths connecting virus-interacting host proteins to "internal-essential" host factors (i.e. the host factors re-verified in the siRNA screen of host factors identified in the Karlas *et al.* screen). The proteins linking internal-essential proteins to virus-interacting proteins are "connecting" candidate proteins for evaluation as host factors of virus replication. The resulting subnetwork contained 1,213 virus-interacting proteins, 38 internal-essential proteins (12 proteins were not in the PPI network), and 1,643 connecting candidate proteins (Table S2 contains the identities and centrality values for all proteins in the subnetwork). As a result of how the subnetwork is constructed, the mean degree of the virus-interacting proteins and the internal-essential proteins were lower than the mean degree of the connecting proteins (see Fig. S2A; ANOVA followed by Tukey post hoc analysis $p < 0.01$). While the degree of connecting proteins does not shift significantly between the total PPI network and the virus subnetwork (Fig. 3A), some proteins with low betweenness have much lower betweenness in the virus subnetwork when compared to the total PPI network (Fig. 3B). Higher betweenness nodes in the total PPI network do not demonstrate dramatic shifts in the virus subnetwork upon comparison. This shift between the total network and virus subnetwork may reveal proteins that are more or less critical to virus replication which cannot be identified in a standard PPI network analysis.

Functional enrichment analysis of the influenza virus-host subnetwork

A functional enrichment analysis was performed using DAVID 6.8's Functional Annotation tool(45). Analysis found that virus-interacting host proteins and connecting (non-internal-

essential) proteins within the virus subnetwork are functionally distinct (see Table 1-2 for abbreviated results, see Table S3 for full results). Gene ontology and pathway analysis found that virus-interacting host proteins are primarily associated with housekeeping and viral replication processes (consistent with the results reported by Watanabe *et al.*(32)), whereas connecting proteins were associated with protein phosphorylation, histone reconfiguration and immune responses. Specifically, the immune response pathways identified are the stimulatory C-type lectin receptor signaling, T-cell receptor signaling, and Fc-epsilon receptor signaling; all of which regulate NFκB activity. Influenza virus is known to manipulate host immune response pathways (specifically NFκB regulating pathways) to promote viral replication(46, 47). These results suggest that the virus subnetwork contains functional information generally unobserved when considering virus-interacting host proteins or internal-essential proteins in isolation.

Connecting proteins of the influenza virus-host subnetwork are more enriched for host factors than are virus-interacting proteins

To evaluate the hypothesis that the "connecting" proteins are likely to be host factors and to simultaneously evaluate if network topology can improve host factor identification, we selected 78 proteins of the subnetwork with the highest (n=39) and lowest betweenness (n=39) and conducted another *in vitro* virus replication assay. HEK293 cells were again transfected with siRNAs targeted to each of the 78 candidate protein's genes and the procedure described previously was performed to determine the proportion of the connecting proteins tested that are host factors of influenza-virus replication. The hit rate is defined as the proportion of proteins tested that significantly down-regulated virus replication.

To evaluate the significance observed in the virus replication screen of the connecting proteins, we compared the observed hit rate to the hit rate observed in a screen the 1,292 virus-interacting host proteins in HEK293 cells (hit rate = 299/1292 = 0.23)(32), in the screen of the 264 host factors from Karlas *et al.* 2010(33) (detailed above), and in a full genome screen for influenza virus host factors in A549 cells (287/22,843 = 0.013)(33). The full genome screen provides the expected hit rate when randomly sampling the PPI. An alternative approach to network-based discovery is to target the nearest neighbors of the virus; a comparison provided by the screen of virus-interacting host proteins. An additional metric is the hit rate observed in our siRNA screen of the host factors identified by Karlas et al (71 out of 264; hit rate = 0.27). Differences between hit rates was compared using the Pearson's chi-squared test when comparing proportions between two binomial groups.

The siRNA screen of the connecting proteins found that connecting proteins were significantly enriched for host factors, but there was no statistically significant advantage in selecting proteins by betweenness (Fig. 4). Of the 78 proteins targeted in the siRNA screen of connecting proteins, a total of 27 significantly reduced virus titers by at least two orders of magnitude; corresponding to 15 categorized as connecting – high betweenness proteins and 12 categorized as connecting – low betweenness proteins. Note that one of the 39 connecting – high betweenness proteins (PLK1) was eliminated from the calculation because both respective siRNAs were cytotoxic (see Table S4). The hit rate of connecting proteins (27/77 = 0.35) was significantly higher than the hit rate observed in the screen of virus-interacting proteins ($p = 0.024$) and in the full genome screen ($p < 2.2*10^{-16}$) but not significantly distinct from the rate observed in the re-screening of the Karlas host factors ($p = 0.21$). When considering the connecting proteins based on their betweenness, the high betweenness had a hit rate of 0.39 (15/38) which was significantly higher than the hit rates

observed in the virus-interacting and full genome screens ($p= 0.032$ and $p < 2.2*10^{-16}$, respectively). High betweenness protein hit rate was higher than the rate observed in the screen of Karlas *et al.* 2010 host factors, but not significantly ($p=0.16$). The low betweenness connecting proteins hit rate was lower than that of the high betweenness connecting proteins (12/39 = 0.31). The difference in hit rates between high and low betweenness proteins was not significant ($p = 0.57$). In all, the screening results suggest that proteins connecting virus-interacting proteins to host factors of influenza virus replication are highly enriched for host factors themselves – significantly more so than proteins which directly interact with virus proteins. However, the topological information from betweenness does not significantly improve host factor identification.

The influenza virus subnetwork is enriched for host factors identified in 6 host factor screens

To determine if host factors identified in previous screens are enriched within the virus subnetwork, we compiled a list of host factors of influenza virus replication identified in at least one of 6 previous screens (33, 48–52) (Table S5). A Fisher exact test for enrichment was used to determine if the connecting proteins or the set of influenza virus-interacting proteins are enriched with host factors identified in these studies relative to the abundance of host factors within the PPI. Both connecting proteins and the virus interacting proteins are significantly enriched for host factors ($p=7.2*10^{-05}$ and $p=1.1*10^{-05}$, respectively; odds ratio = 1.4 and 1.5, respectively). When compared, there is no significant difference in the enrichment of host factors between connecting proteins and virus-interacting proteins ($p = 0.48$; odds ratio = 0.92). To ensure the host factors identified in the Karlas et al study were not creating bias in the enrichment result, the enrichment analysis was repeated using host factors identified in all studies except the Karlas study. Again, connecting proteins and virus

interacting proteins are significantly enriched for host factors ($p=1.8*10^{-06}$ and $p=3.2*10^{-03}$, respectively; odds ratio = 1.5 and 1.34, respectively) and no significant difference in the enrichment of host factors between connecting proteins and virus-interacting proteins was found ($p = 0.49$).

DISCUSSION

Network approaches have demonstrated their potential impact on health related research including gene/protein characterization and drug design and side effects (14, 18, 19, 22, 36, 53) yet demonstrations that network information can inform drug target discovery is still limited. Here, we present the first confirmation that virus and host protein interaction data can be integrated to improve large-scale drug target discovery (specifically antiviral target discovery) and reveal additional insights into virus-host interactions. The position of virus-interacting proteins suggest that the influenza virus has evolved to interact with proteins that heavily influence network behavior. Additionally, virus-interacting proteins are closely clustered in the network. This may be a result of attempts by the virus to manipulate specific biological functions (as proteins with shared biological functions tend to cluster in PPI networks(54)) signifying that influenza virus has parallel available pathways to engage with host biological functions. Previous studies have found that host factors of virus replication (not necessarily virus-interacting host proteins) have also been observed to cluster within the PPI network(41). Further analysis on network clustering host factors of interest is needed to determine if these two observations are independent of one another.

The observation that host-virus interaction data can be leveraged to improve virus replication host factor discovery is unlikely to be affected by off-target concerns associated with siRNA screens. Off-target concerns often challenge siRNA studies though changes to

experimental protocols (such as requiring multiple siRNA hits per targeted gene or changing siRNA concentrations) can only moderately improve false positive rates (55–57). The protocol used in this study was not optimal to ensure the characterization of any one targeted gene. As such, the hit rates of gene groups are compared. Protocols between these experiments and those used for comparisons are either identical (32) or very similar (33), suggesting that off-target rates across the tested groups are unlikely to explain the differences in observed hit rates.

The variability and incompleteness of PPI data as well as the limited agreement between influenza virus replication screens are well known concerns for network-based drug target discovery. The possibility that virus-host interaction data is skewed towards well studied networks could also have an effect on the clustering result in virus-interacting proteins. However, the enrichment of host proteins important for influenza virus replication within the constructed virus subnetwork demonstrates that even with these possible shortcomings, PPI network analyses have the power to identify important host factors for influenza virus replication. The antiviral protein candidate screen performed in this study further supports the use of PPI data during candidate prioritization with significant hit rates against virus-interacting proteins and randomly targeted proteins.

The observation that betweenness does not significantly improve host factor prediction suggests that alternative topology measures should be considered. There were several reasons why betweenness was selected. Biological pathways are known to have several alternative routes to maintaining cellular operations; a key feature of biological robustness (58–60). Biological networks are also theorized to have a bow tie-like structure that suggests a natural bottlenecking within the PPI near critical cellular machinery(61). These concepts

together suggest targeting bottlenecks (high betweenness proteins) as a means of mitigating escape via alternative paths. It was also the concern of alternative pathways as to why the set of virus interacting proteins was not limited to confirmed host factors of influenza virus replication. In future work, other network topology measures, e.g. degree, Burt's Constraint, or closeness, could be tested in the subnetwork and subnetwork construction could be varied to consider different subsets of either the virus interacting proteins or the internal host factors. Even so, the results suggest that the construction of the virus-specific subnetwork provides major advantages in host factor discovery and can significantly expand drug candidate repertoires beyond virus-interacting proteins. Furthermore, since the connecting proteins do not directly interact with the virus, they may be more resistant concerns related to drug-mediated selective pressure.

Another interesting continuation of this study would identify the cause of connecting proteins' effect on virus replication. The mechanism by which each host factor is regulating virus replication may offer additional clues for drug candidate prioritization efforts. Overall, this PPI-based study provides insight into the network characteristics of virus-host interactions and supports the idea that the influenza virus evolved to interact with host proteins in dominant network positions in order to maximally manipulate host cells.

METHODS AND MATERIALS

Protein-protein interaction network construction and analysis

Protein-protein interaction data was downloaded from the Human Integrated Protein-Protein Interaction rEference (HIPPIE) database(37) (version 1.4). Interactions with a confidence score less than 0.7 were removed. The interaction data was then analyzed with the igraph package in R. The interaction data resulted in one large network containing 9,969

377 nodes and 86 smaller disconnected networks (most with 2 nodes, all contained 7 or fewer)
378 which were removed from the study. The final human PPI contained 9,969 proteins and
379 57,615 interactions.

380 All PPI topology analyses were performed with the R library igraph version 1.0.1(62).

381 *Statistical analyses and graphics packages*

382 All statistical tests were performed in R 3.2.2 using the functions prop.test, fisher.test,
383 pairwise.t.test or wilcoxon.test (which performs a Mann-Whitney-U test) as appropriate.
384 Prop.test and fisher.test both compare outcome proportions between binomial groups with
385 the latter being more precise for small group sizes. Graphics were produced with either the
386 default graphing features of R or with the ggplot2 library (63).

387 *Cells and Viruses*

388 Influenza A/WSN/ 33 virus (WSN; H1N1) was generated using reverse genetics(64). HEK293
389 cells were cultured in DMEM (Sigma-Aldrich) supplemented with 10% FCS (10% FCS/DMEM)
390 and antibiotics at 37°C in 5% CO_2. Virus plaque titers were determined by plaque assay in
391 Madin–Darby canine kidney (MDCK) cells. MDCK cells were cultured in Eagle's MEM
392 (GIBCO) with 5% NCS at 37°C in 5% CO_2.

393 *siRNA Treatment*

394 siRNA treatment procedure, cell viability and virus titer determination are described in
395 detail in Watanabe et al 2014. Briefly, two siRNAs per candidate gene were selected from a
396 predesigned genome-wide human siRNA library (FlexTube siRNA; QIAGNE). AllStars
397 Negative Control siRNA (QIAGEN) was served as a negative control. The siRNA against the
398 NP gene of WSN virus (GGA UCU UAU UUC UUC GGA GUU) purchased from Sigma-Aldrich
399 was used as a positive control. HEK293 cells were transfected twice with 25 nM (final

concentration, 50 nM) of siRNA duplexes using RNAiMAX (Invitrogen). At 24 hr after the second transfection, cell viability was determined using the CellTiter-Glo assay system (Promega) following manufacturer's instructions. To assess influenza virus replication, two parallel sets of siRNA-transfected cells were infected with 50 plaque forming units (pfu) of WSN virus per well of a 24-well tissue culture plate at 24 hr after the second siRNA transfection. At 48 hr post-infection, supernatants were harvested and virus titers determined by plaque assay in MDCK cells.

Quantitative reverse transcription-PCR

To confirm the down-regulation of host genes by their respective target siRNAs, quantitative reverse transcription-PCR (qRT-PCR) experiments were performed. Table S6 provides a complete list of primer sequences. HEK 293 cells, transfected twice with 25 nM of siRNA (final concentration, 50 nM), were lysed at 48 h post-transfection and total RNA was extracted by using the Maxwell 16 LEV simplyRNA Tissue Kit (Promega). Reverse transcription was performed by using ReverTra Ace qPCR RT Master Mix (TOYOBO, Osaka, Japan) or SuperScript III Reverse Transcriptase (Invitrogen). The synthesized cDNA was subjected to quantitative PCR with primers specific for each gene by using the THUNDERBIRD SYBR qPCR Mix (TOYOBO). The relative mRNA expression levels of each gene were calculated by the $\Delta\Delta$Ct method using beta-actin as internal control. Primer sequences are available upon request.

Determining candidate proteins involved in influenza virus replication

For each set of siRNAs, the virus titers from cells treated with siRNAs were normalized by the titers obtained from cell treated with AllStars Negative Control siRNA (Table S7). siRNAs that reduced cell viability by more than 40% relative to that of AllStars Negative Control

siRNA-treated cells were not considered for further analysis. Unlike our previous study(32), LOESS regression was not needed (Fig. S3). A two-sided, unpaired Student's t test was used to compare the mean virus titers in cells treated with gene-specific siRNAs with those in cells treated with AllStars Negative Control siRNA. Holm's method for multiple comparisons was then applied to the p values.

ACKNOWLEDGEMENTS

We thank Naomi Fujimoto, Tomoko Kuwahara and Kazue Goto for technical assistance.

DECLARATION OF INTEREST

The authors declare no competing interests.

REFERENCES

1. van Hoek AJ, Underwood A, Jit M, Miller E, Edmunds WJ. 2011. The impact of pandemic influenza H1N1 on health-related quality of life: A prospective population-based study. PLoS One 6:1–6.

2. Perdue ML, Bright RA. 2011. United States of America Department of Health and Human Services support for advancing influenza vaccine manufacturing in the developing world. Vaccine 29:A48–A50.

3. Davis MM, Wortley PM, Ndiaye SM, Woods MG, Clark SJ. 2004. National availability of influenza vaccine among medical subspecialty practices. Am J Prev Med 26:307–310.

4. Treanor J. 2004. Influenza Vaccine — Outmaneuvering Antigenic Shift and Drift. N Engl J Med 350:218–220.

444 5. Neumann G, Chen H, Gao GF, Shu Y, Kawaoka Y. 2010. H5N1 influenza viruses:
445 outbreaks and biological properties. Cell Res 20:51–61.

446 6. Gao R, Cao B, Hu Y, Feng Z, Wang D, Hu W, Chen J, Jie Z, Qiu H, Xu K, Xu X, Lu H, Zhu
447 W, Gao Z, Xiang N, Shen Y, He Z, Gu Y, Zhang Z, Yang Y, Zhao X, Zhou L, Li XX, Zou S,
448 Zhang YY, Yang L, Guo J, Dong J, Li Q, Dong L, Zhu Y, Bai T, Wang S, Hao P, Yang W,
449 Han J, Yu H, Li D, Gao GF, Wu G, Wang Y, Yuan Z, Shu Y. 2013. Human infection with a
450 novel avian-origin influenza A (H7N9) virus. N Engl J Med 368:1888–97.

451 7. Li Q, Zhou L, Zhou M, Chen Z, Li F, Wu H, Xiang N, Chen E, Tang F, Wang D, Meng L,
452 Hong Z, Tu W, Cao Y, Li L, Ding F, Liu B, Wang M, Xie R, Gao R, Li X, Bai T, Zou S, He J,
453 Hu J, Xu Y, Chai C, Wang S, Gao Y, Jin L, Zhang Y, Luo H, Yu H, He J, Li Q, Wang X, Gao
454 L, Pang X, Liu G, Yan Y, Yuan H, Shu Y, Yang W, Wang Y, Wu F, Uyeki TM, Feng Z. 2014.
455 Epidemiology of human infections with avian influenza A(H7N9) virus in China. N Engl
456 J Med 370:520–32.

457 8. Yen H-L, Webster RG. 2009. Vaccines for Pandemic Influenza, p. 3–24. *In* Compans,
458 WR, Orenstein, AW (eds.), . Springer Berlin Heidelberg, Berlin, Heidelberg.

459 9. Davies WL, Grunert RR, Haff RF, McGahen JW, Neumayer EM, Paulshock M, Watts JC,
460 Wood TR, Hermann EC, Hoffmann CE. 1964. Antiviral Activity of 1-Adamantanamine
461 (Amantadine). Science (80-) 144:862–863.

462 10. Hayden FG. 2001. Perspectives on antiviral use during pandemic influenza. Philos
463 Trans R Soc Lond B Biol Sci 356:1877–1884.

464 11. Tamura D, Mitamura K, Yamazaki M, Fujino M, Nirasawa M, Kimura K, Kiso M,
465 Shimizu H, Kawakami C, Hiroi S, Takahashi K, Hata M, Minagawa H, Kimura Y, Kaneda

S, Sugita S, Horimoto T, Sugaya N, Kawaoka Y. 2009. Oseltamivir-resistant influenza a viruses circulating in Japan. J Clin Microbiol 47:1424–7.

12. Bright R a, Shay DK, Shu B, Cox NJ, Klimov AI. 2006. Adamantane resistance among influenza A viruses isolated early during the 2005-2006 influenza season in the United States. JAMA 295:891–894.

13. Barabasi A-L, Gulbahce N, Loscalzo J. 2011. Network medicine: a network-based approach to human disease. Nat Rev Genet 12:56–68.

14. Kotlyar M, Fortney K, Jurisica I. 2012. Network-based characterization of drug-regulated genes, drug targets, and toxicity. Methods 57:499–507.

15. Korth MJ, Tchitchek N, Benecke AG, Katze MG. 2012. Systems approaches to influenza-virus host interactions and the pathogenesis of highly virulent and pandemic viruses. Semin Immunol 1–12.

16. Liu YY, Slotine JJ, Barabási AL. 2011. Controllability of complex networks. Nature 473:167–173.

17. Burt RS. 2004. Structural Holes and Good Ideas. Am J Sociol 110:349–399.

18. Jonsson PF, Bates PA. 2006. Global topological features of cancer proteins in the human interactome. Bioinformatics 22:2291–2297.

19. Hase T, Tanaka H, Suzuki Y, Nakagawa S, Kitano H. 2009. Structure of protein interaction networks and their implications on drug design. PLoS Comput Biol 5.

20. Yildirim M a, Goh K-I, Cusick ME, Barabási A-L, Vidal M. 2007. Drug-target network. Nat Biotechnol 25:1119–26.

487 21. Zhu M, Gao L, Li X, Liu Z. 2009. Identifying drug-target proteins based on network
488 features. Sci China C Life Sci 52:398–404.

489 22. Lopes TJS, Shoemaker JE, Matsuoka Y, Kawaoka Y, Kitano H. 2015. Identifying
490 problematic drugs based on the characteristics of their targets. Front Pharmacol 6.

491 23. Jin S, Li Y, Pan R, Zou X. 2014. Characterizing and controlling the inflammatory
492 network during influenza A virus infection. Sci Rep 4:3799.

493 24. Schaefer MH, Lopes TJS, Mah N, Shoemaker JE, Matsuoka Y, Fontaine J-F, Louis-Jeune
494 C, Eisfeld AJ, Neumann G, Perez-Iratxeta C, Kawaoka Y, Kitano H, Andrade-Navarro M
495 a. 2013. Adding protein context to the human protein-protein interaction network to
496 reveal meaningful interactions. PLoS Comput Biol 9:e1002860.

497 25. Shoemaker JE, Fukuyama S, Eisfeld AJ, Muramoto Y, Watanabe S, Watanabe T,
498 Matsuoka Y, Kitano H, Kawaoka Y. 2012. Integrated network analysis reveals a novel
499 role for the cell cycle in 2009 pandemic influenza virus-induced inflammation in
500 macaque lungs. BMC Syst Biol 6:117.

501 26. Taye B, Vaz C, Tanavde V, Kuznetsov VA, Eisenhaber F, Sugrue RJ, Maurer-Stroh S.
502 2017. Benchmarking selected computational gene network growing tools in context
503 of virus-host interactions. Sci Rep 7:5805.

504 27. Heaton NS, Moshkina N, Fenouil R, Gardner TJ, Aguirre S, Shah PS, Zhao N,
505 Manganaro L, Hultquist JF, Noel J, Sachs DH, Hamilton J, Leon PE, Chawdury A,
506 Tripathi S, Melegari C, Campisi L, Hai R, Metreveli G, Gamarnik A V., García-Sastre A,
507 Greenbaum B, Simon V, Fernandez-Sesma A, Krogan NJ, Mulder LCF, van Bakel H,
508 Tortorella D, Taunton J, Palese P, Marazzi I. 2016. Targeting Viral Proteostasis Limits

509 Influenza Virus, HIV, and Dengue Virus Infection. Immunity 44:46–58.

510 28. Meyniel-Schicklin L, de Chassey B, André P, Lotteau V. 2012. Viruses and Interactomes
511 in Translation. Mol Cell Proteomics 11:M111.014738.

512 29. Germain M-A, Chatel-Chaix L, Gagné B, Bonneil É, Thibault P, Pradezynski F, de
513 Chassey B, Meyniel-Schicklin L, Lotteau V, Baril M, Lamarre D. 2014. Elucidating novel
514 hepatitis C virus-host interactions using combined mass spectrometry and functional
515 genomics approaches. Mol Cell Proteomics 13:184–203.

516 30. Wang L, Fu B, Li W, Patil G, Liu L, Dorf ME, Li S. 2017. Comparative influenza protein
517 interactomes identify the role of plakophilin 2 in virus restriction. Nat Commun 8.

518 31. Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM,
519 Dejesus P, Che J, Mulder LCF, Yángüez E, Andenmatten D, Pache L, Manicassamy B,
520 Albrecht RA, Gonzalez MG, Nguyen Q, Brass A, Elledge S, White M, Shapira S,
521 Hacohen N, Karlas A, Meyer TF, Shales M, Gatorano A, Johnson JR, Jang G, Johnson T,
522 Verschueren E, Sanders D, Krogan N, Shaw M, König R, Stertz S, García-Sastre A,
523 Chanda SK. 2015. Meta- and Orthogonal Integration of Influenza "oMICs" Data
524 Defines a Role for UBR4 in Virus Budding. Cell Host Microbe 18:723–735.

525 32. Watanabe T, Kawakami E, Shoemaker JE, Lopes TJS, Matsuoka Y, Tomita Y, Kozuka-
526 Hata H, Gorai T, Kuwahara T, Takeda E, Nagata A, Takano R, Kiso M, Yamashita M,
527 Sakai-Tagawa Y, Katsura H, Nonaka N, Fujii H, Fujii K, Sugita Y, Noda T, Goto H,
528 Fukuyama S, Watanabe S, Neumann G, Oyama M, Kitano H, Kawaoka Y. 2014.
529 Influenza virus-host interactome screen as a platform for antiviral drug development.
530 Cell Host Microbe 16:795–805.

531 33. Karlas A, Machuy N, Shin Y, Pleissner K-P, Artarini A, Heuer D, Becker D, Khalil H,
532 Ogilvie LA, Hess S, Mäurer AP, Müller E, Wolff T, Rudel T, Meyer TF. 2010. Genome-
533 wide RNAi screen identifies human host factors crucial for influenza virus replication.
534 Nature 463:818–22.

535 34. He X, Zhang J. 2006. Why do hubs tend to be essential in protein networks? PLoS
536 Genet 2:0826–0834.

537 35. Zhao J, Yang T-H, Huang Y, Holme P. 2011. Ranking candidate disease genes from
538 gene expression and protein interaction: a Katz-centrality based approach. PLoS One
539 6:e24306.

540 36. Wang X, Thijssen B, Yu H. 2013. Target essentiality and centrality characterize drug
541 side effects. PLoS Comput Biol 9:e1003119.

542 37. Schaefer MH, Fontaine JF, Vinayagam A, Porras P, Wanker EE, Andrade-Navarro MA.
543 2012. Hippie: Integrating protein interaction networks with experiment based quality
544 scores. PLoS One 7:1–8.

545 38. Freeman LC. 1977. A Set of Measures of Centrality Based on Betweenness.
546 Sociometry 40:35–41.

547 39. de Chassey B, Navratil V, Tafforeau L, Hiet MS, Aublin-Gex A, Agaugué S, Meiffren G,
548 Pradezynski F, Faria BF, Chantier T, Le Breton M, Pellet J, Davoust N, Mangeot PE,
549 Chaboud A, Penin F, Jacob Y, Vidalain PO, Vidal M, André P, Rabourdin-Combe C,
550 Lotteau V. 2008. Hepatitis C virus infection protein network. Mol Syst Biol 4:1–12.

551 40. Khadka S, Vangeloff AD, Zhang C, Siddavatam P, Heaton NS, Wang L, Sengupta R,
552 Sahasrabudhe S, Randall G, Gribskov M, Kuhn RJ, Perera R, LaCount DJ. 2011. A

553 Physical Interaction Network of Dengue Virus and Human Proteins. Mol Cell
554 Proteomics 10:M111.012187.

555 41. Hao L, He Q, Wang Z, Craven M, Newton M a, Ahlquist P. 2013. Limited agreement of
556 independent RNAi screens for virus-required host genes owes more to false-negative
557 than false-positive factors. PLoS Comput Biol 9:e1003235.

558 42. Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B,
559 Vitols E, Vijayadamodar G, Pochart P, Machineni H, Welsh M, Kong Y, Zerhusen B,
560 Malcolm R, Varrone Z, Collis A, Minto M, Burgess S, McDaniel L, Stimpson E, Spriggs F,
561 Williams J, Neurath K, Ioime N, Agee M, Voss E, Furtak K, Renzulli R, Aanensen N,
562 Carrolla S, Bickelhaupt E, Lazovatsky Y, DaSilva A, Zhong J, Stanyon CA, Finley RL,
563 White KP, Braverman M, Jarvie T, Gold S, Leach M, Knight J, Shimkets RA, McKenna
564 MP, Chant J, Rothberg JM. 2003. A protein interaction map of Drosophila
565 melanogaster. Science (80-) 302:1727–1736.

566 43. Andrews T, Honti F, Pfundt R, Leeuw N de, Hehir-Kwa J, Silfhout AV, Vries B de,
567 Webber C. 2015. The clustering of functionally related genes contributes to CNV-
568 mediated disease. Genome Res 25:802–813.

569 44. Ruepp A, Brauner B, Dunger-Kaltenbach I, Frishman G, Montrone C, Stransky M,
570 Waegele B, Schmidt T, Doudieu ON, Stümpflen V, Mewes HW. 2008. CORUM: the
571 comprehensive resource of mammalian protein complexes. Nucleic Acids Res
572 36:D646-50.

573 45. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW,
574 Lane HC, Lempicki R a. 2007. DAVID Bioinformatics Resources: expanded annotation

575 database and novel algorithms to better extract biology from large gene lists. Nucleic
576 Acids Res 35:W169-75.

577 46. Hale BG, Randall RE, Ortín J, Jackson D. 2008. The multifunctional NS1 protein of
578 influenza A viruses. J Gen Virol 89:2359–76.

579 47. Pauli E-K, Schmolke M, Wolff T, Viemann D, Roth J, Bode JG, Ludwig S. 2008. Influenza
580 A Virus Inhibits Type I IFN Signaling via NF-κB-Dependent Induction of SOCS-3
581 Expression. PLoS Pathog 4:e1000196.

582 48. Hao L, Sakurai A, Watanabe T, Sorensen E, Nidom CA, Newton MA, Ahlquist P,
583 Kawaoka Y. 2008. Drosophila RNAi screen identifies host genes important for
584 influenza virus replication. Nature 454:890–893.

585 49. Shapira SD, Gat-Viks I, Shum BO, Dricot A, de Grace MM, Wu L, Gupta PB, Hao T,
586 Silver SJ, Root DE, Hill DE, Regev A, Hacohen N. 2009. A physical and regulatory map
587 of host-influenza interactions reveals pathways in H1N1 infection. Cell 139:1255–
588 1267.

589 50. Brass AL, Huang IC, Benita Y, John SP, Krishnan MN, Feeley EM, Ryan BJ, Weyer JL, van
590 der Weyden L, Fikrig E, Adams DJ, Xavier RJ, Farzan M, Elledge SJ. 2009. The IFITM
591 proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and
592 dengue virus. Cell 139:1243–1254.

593 51. Sui B, Bamba D, Weng K, Ung H, Chang S, Van Dyke J, Goldblatt M, Duan R, Kinch MS,
594 Li WB. 2009. The use of Random Homozygous Gene Perturbation to identify novel
595 host-oriented targets for influenza. Virology 387:473–481.

596 52. König R, Stertz S, Zhou Y, Inoue A, Hoffmann H-H, Bhattacharyya S, Alamares JG,

Tscherne DM, Ortigoza MB, Liang Y, Gao Q, Andrews SE, Bandyopadhyay S, De Jesus P, Tu BP, Pache L, Shih C, Orth A, Bonamy G, Miraglia L, Ideker T, García-Sastre A, Young JAT, Palese P, Shaw ML, Chanda SK. 2010. Human host factors required for influenza virus replication. Nature 463:813–7.

53. Xia J, Sun J, Jia P, Zhao Z. 2011. Do cancer proteins really interact strongly in the human protein-protein interaction network? Comput Biol Chem 35:121–125.

54. Poyatos JF, Hurst LD. 2006. Is optimal gene order impossible? Trends Genet 22:420–423.

55. Echeverri CJ, Beachy PA, Baum B, Boutros M, Buchholz F, Chanda SK, Downward J, Ellenberg J, Fraser AG, Hacohen N, Hahn WC, Jackson AL, Kiger A, Linsley PS, Lum L, Ma Y, Mathey-Prévôt B, Root DE, Sabatini DM, Taipale J, Perrimon N, Bernards R. 2006. Minimizing the risk of reporting false positives in large-scale RNAi screens. Nat Methods 3:777–779.

56. Zhang XD. 2010. An effective method for controlling false discovery and false nondiscovery rates in genome-scale RNAi screens. J Biomol Screen 15:1116–1122.

57. Birmingham A, Selfors LM, Forster T, Wrobel D, Kennedy CJ, Shanks E, Santoyo-Lopez J, Dunican DJ, Long A, Kelleher D, Smith Q, Beijersbergen RL, Ghazal P, Shamu CE. 2009. Statistical methods for analysis of high-throughput RNA interference screens. Nat Methods 6:569–75.

58. Gong Y, Zhang Z. 2005. Alternative signaling pathways: When, where and why? FEBS Lett 579:5265–5274.

59. Weng G, Bhalla US, Iyengar R. 1999. Complexity in biological signaling systems.

619 Science 284:92–6.

620 60. Kitano H. 2004. Biological robustness. Nat Rev Genet 5:826–837.

621 61. Csete M, Doyle J. 2004. Bow ties, metabolism and disease. Trends Biotechnol 22:446–
622 50.

623 62. Gabor Csardi and Tamas Nepusz. 2006. The igraph software package for complex
624 network research. InterJournal Complex Sy.

625 63. Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New
626 York.

627 64. Neumann G, Watanabe T, Ito H, Watanabe S, Goto H, Gao P, Hughes M, Perez DR,
628 Donis R, Hoffmann E, Hobom G, Kawaoka Y. 1999. Generation of influenza A viruses
629 entirely from cloned cDNAs. Proc Natl Acad Sci U S A 96:9345–50.

630

631 ## FIGURES

632 **Figure 1 The virus-interacting network and the virus subnetwork.** (a) The virus-interacting
633 network is created from human host-PPI data combined with virus-host protein interaction
634 data. (b) The virus subnetwork was isolated from the complete human PPI network by
635 collecting all interactions involved in the shortest paths (red) that connect influenza virus-
636 interacting proteins (blue) to human proteins essential to virus replication (e.g. the internal-
637 essential proteins; colored orange). The connecting proteins (colored black) are candidates
638 to be evaluated for their antiviral properties.

639 **Figure 2 The network topological characteristics of virus-interacting host proteins.** The
640 distributions of the (A) degree and (B) betweenness of virus-interacting proteins and all
641 proteins in the human PPI network. An $\varepsilon = 0.01$ was added to the betweenness to facilitate
642 log scaling. The cumulative distributions (thick, red lines) of the shortest distances
643 connecting host proteins in the PPI network that interact with viral (C) PB1, (D) HA, (E) NS1
644 proteins or (F) the set of all viral proteins. As a control, the cumulative distribution of
645 distances was iteratively determined (N=100) by randomly sampled host proteins in the PPI
646 network (thin, black lines). The number of proteins sampled on each iteration was equal to
647 the number of interacting host proteins of each virus protein (or set of viral proteins).

648 **Figure 3 Network characteristics of the virus subnetwork.** Panels (A) and (B) compare the
649 degree and betweenness, respectively, of the connecting proteins in the whole PPI network
650 and the virus subnetwork.

651 **Table 1 Functional Enrichment Analysis of Virus Subnetwork.** Functional enrichment
652 analysis of connecting proteins within the virus subnetwork. Proteins were analyzed using
653 DAVID.

654 **Table 2 Functional Enrichment Analysis of Virus Subnetwork.** Functional enrichment
655 analysis of virus-interacting proteins within the virus subnetwork. Proteins were analyzed
656 using DAVID.

657 **Figure 4 Comparison of hit rates.** The hit rates are reported for all tested connecting
658 proteins and connecting proteins with high or low betweenness in the virus subnetwork.
659 These hit rates are compared to hit rates observed from a previous screen of virus-
660 interacting host proteins (labeled "Virus-Interacting Proteins") [32], from applying our
661 screening methodology to host factors identified in a screen by Karlas *et al*. (labeled "Karlas

host factors") and from a genome-wide screen [33]. Prop.test in R was used to determine the significance of the difference in hit rates observed for binomial groups. * indicates a $p < 0.05$ and ** indicates a $p < 0.01$.

SUPPLEMENTAL FIGURES

Fig S1 The distributions of the (a) degree and (b) betweenness of the interaction partners of each of the 11 virus proteins. The y axis lists the particular virus protein, and the x axis demonstrates distributions of the centrality measures of the virus protein's interaction partners within the human PPI network. The distributions for all proteins in the human PPI network (labeled "All") and the set of proteins that interacted with any of the virus proteins ("VB") are included for comparison.

Fig S2 Boxplot of the degree and betweenness distributions for connecting (candidate) proteins, virus-interacting proteins, and internal essential proteins. Black lines indicate the median for each population.

Fig S3 The mean log fold change (LFC) vs the mean fold change (FC) in cell viability for all 156 gene-specific siRNAs tested. Cyan and green points highlight data corresponding to the 24 negative and positive control siRNAs (i.e., AllStars Negative Control siRNA and 25 siRNA against influenza virus NP gene, respectively). The broken ride line is the LOESS regression curve, showing that virus growth was not dependent on cell viability.

Table S1 Effects of siRNAs targeting host factors identified to be important for influenza virus replication by Karlas et al. (Nature, 2010) on virus production. Note that two siRNA's

were used per Entrez Gene ID. Sheet 2, labeled “untested host factors”, lists host factors that were identified in the Karlas screen but were not evaluated in this study.

Table S2 The degree and betweenness of proteins in virus-host interaction subnetwork. The symbol, description and Entrez Gene ID of each protein are provided in the first three columns. Proteins tagged with a 1 in the "Virus-interacting" and "Internal-Essential" columns identify proteins were associated with a virus protein in the co-immunoprecipitation study or identified as essential but not directly associated with a virus protein, respectively. The last three columns provide the protein's degree and betweenness in the subnetwork and identify which proteins were selected for further testing.

Table S3 DAVID Functional Annotation Tool results for virus-interacting proteins and connecting proteins of the influenza virus subnetwork. Full results include the clustering, chart, and table outputs from DAVID 6.8.

Table S4 Effects of siRNAs Targeting Host Factors with High or Low Betweenness in the Virus-Host Subnetwork on Virus Production.

Table S5 Hit-lists of genes identified in 6 independent genome-wide screens. Studies include König et al (2010), Brass et al (2009), Shapira et al (2009), Hao et al (2008), Karlas et al (2010), and Sui et al (2009).

Table S6 A list of primers used for qPCR.

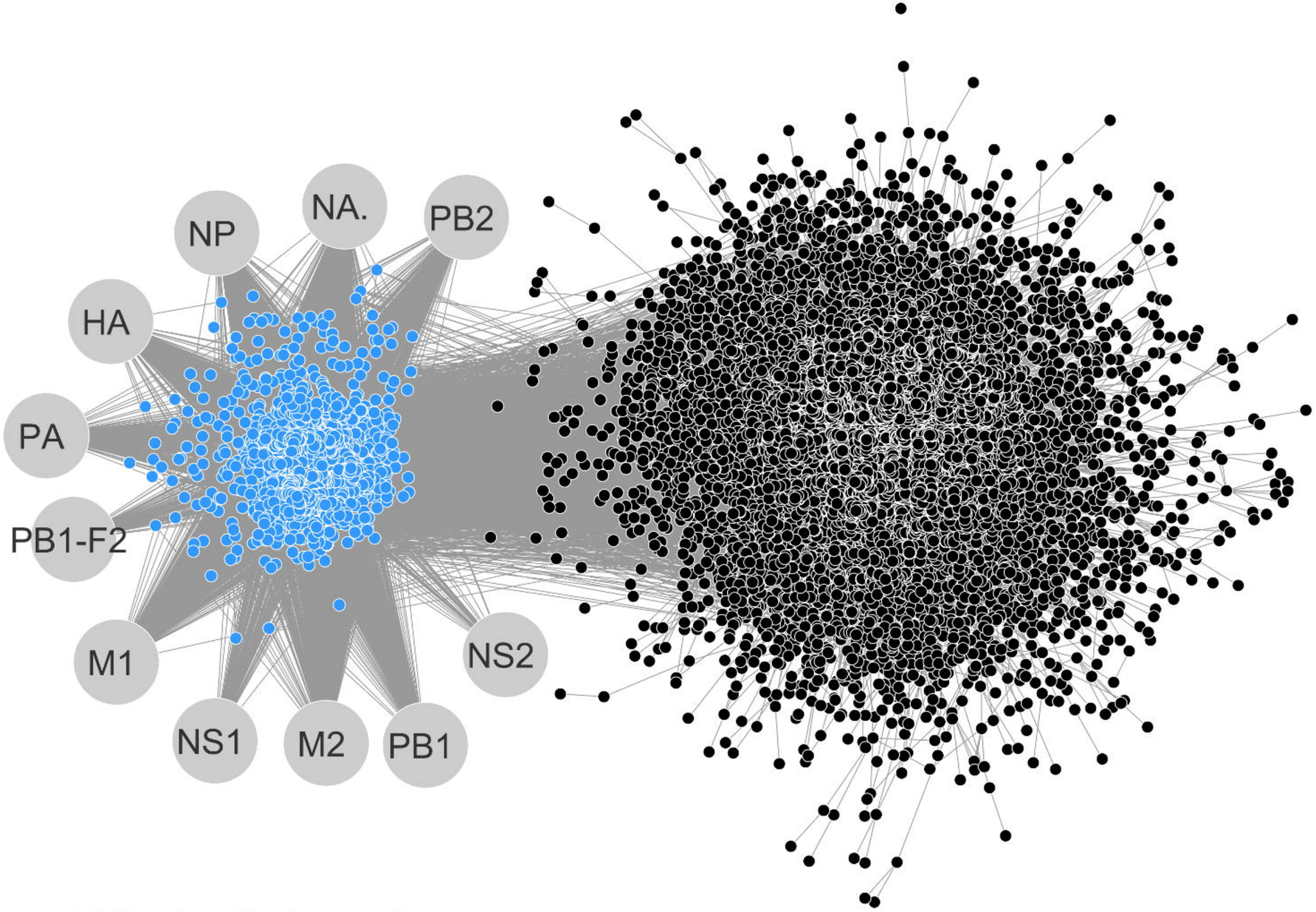
Table S7 Virus titers observed in HEK293 cells.

A

Virus protein

Virus interacting host protein

Host protein



B

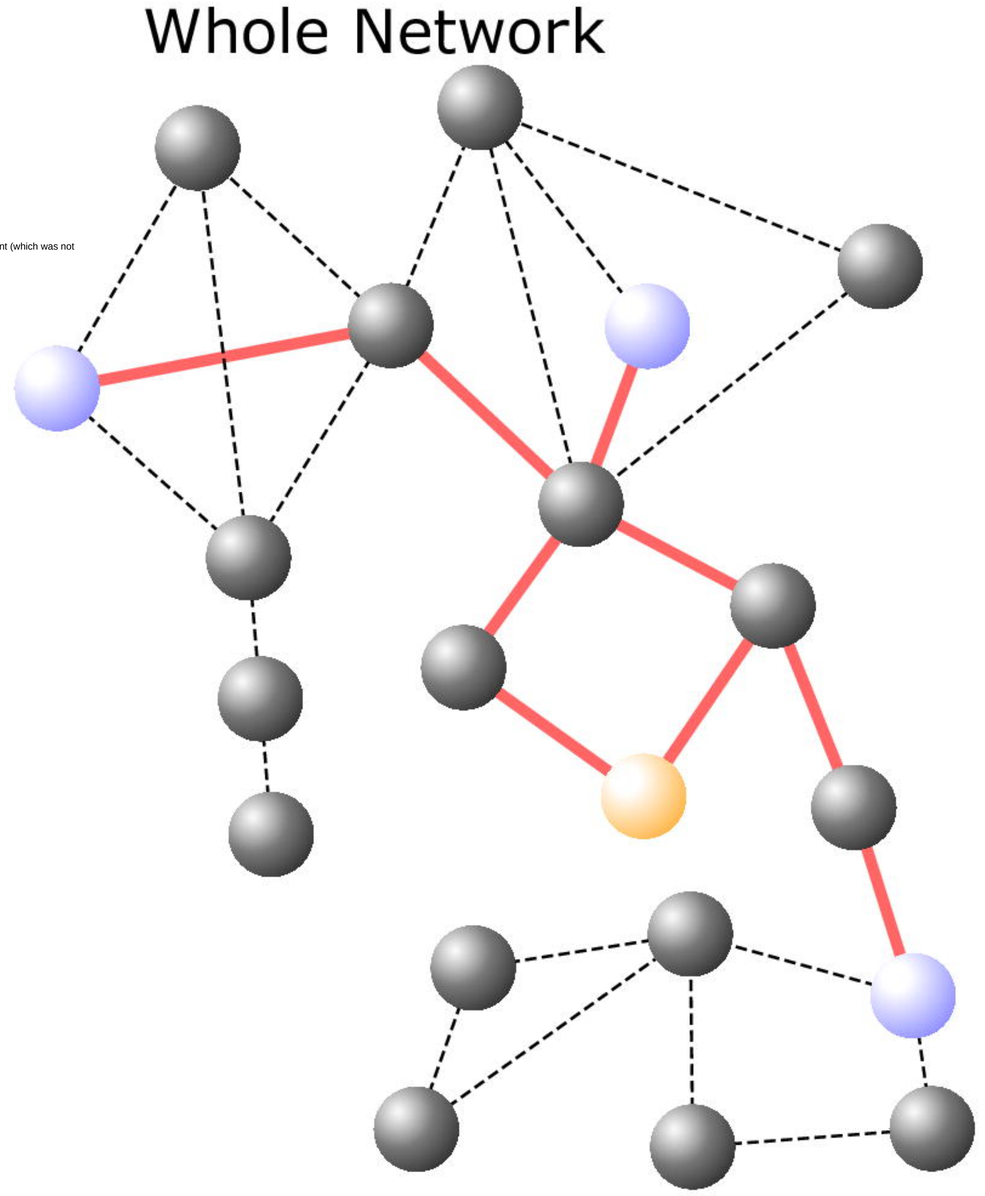
Virus interacting host protein

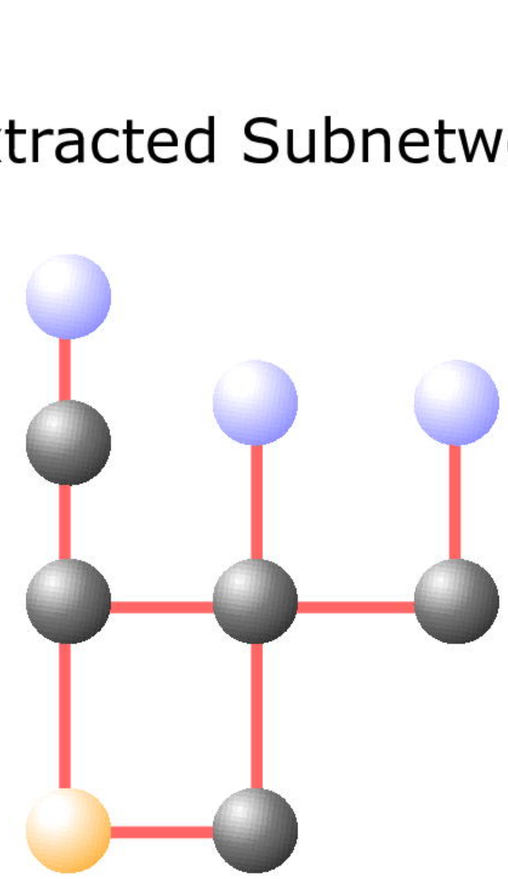
Internal-essential host protein

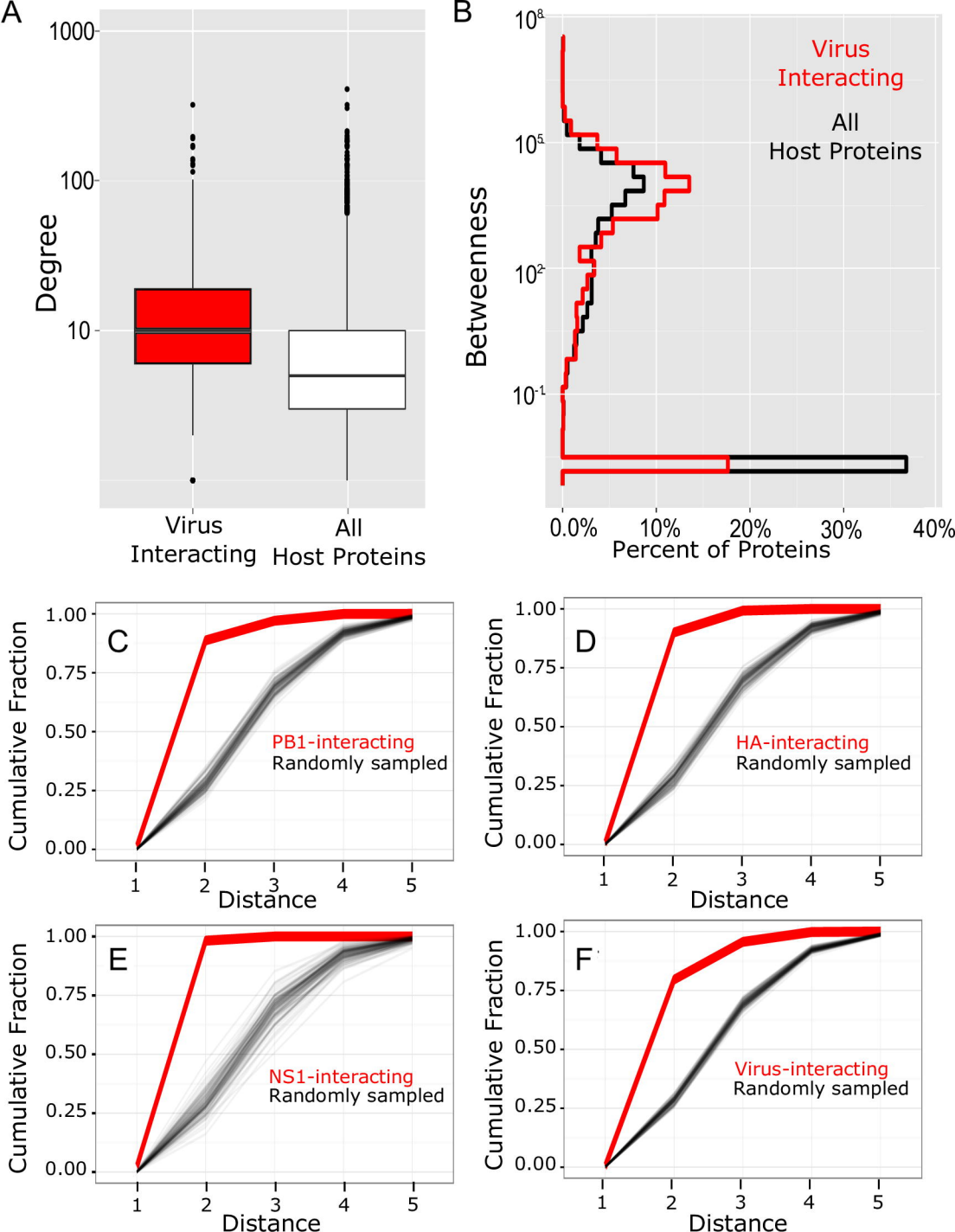
Connecting Host Protein

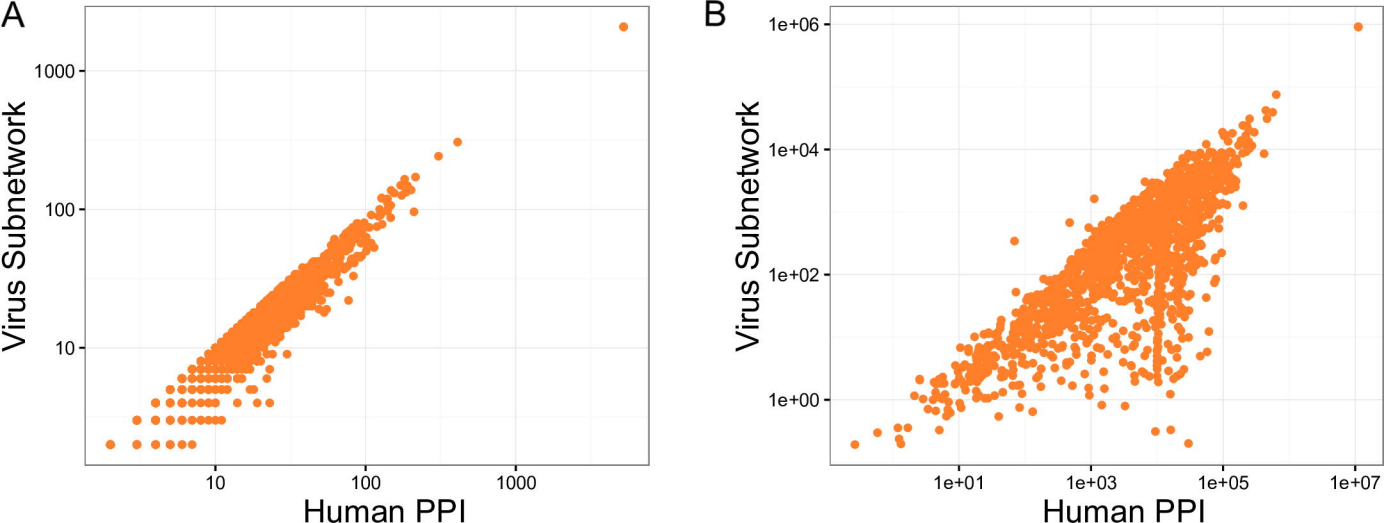
---------- Interaction

Shortest path






A
1000
100
10
Degree
Virus Interacting
All Host Proteins
B
10^{8}
10^{5}
10^{2}
10^{-1}
Betweenness
Virus Interacting
All Host Proteins
0.0%
10%
20%
30%
40%
Percent of Proteins
C
PB1-interacting
Randomly sampled
D
HA-interacting
Randomly sampled
E
NS1-interacting
Randomly sampled
F
Virus-interacting
Randomly sampled
Cumulative Fraction
1.00
0.75
0.50
0.25
0.00
1
2
3
4
5
Distance


A
Virus Subnetwork
1000
100
10
10
100
1000
Human PPI
B
Virus Subnetwork
1e+06
1e+04
1e+02
1e+00
1e+01
1e+03
1e+05
1e+07
Human PPI

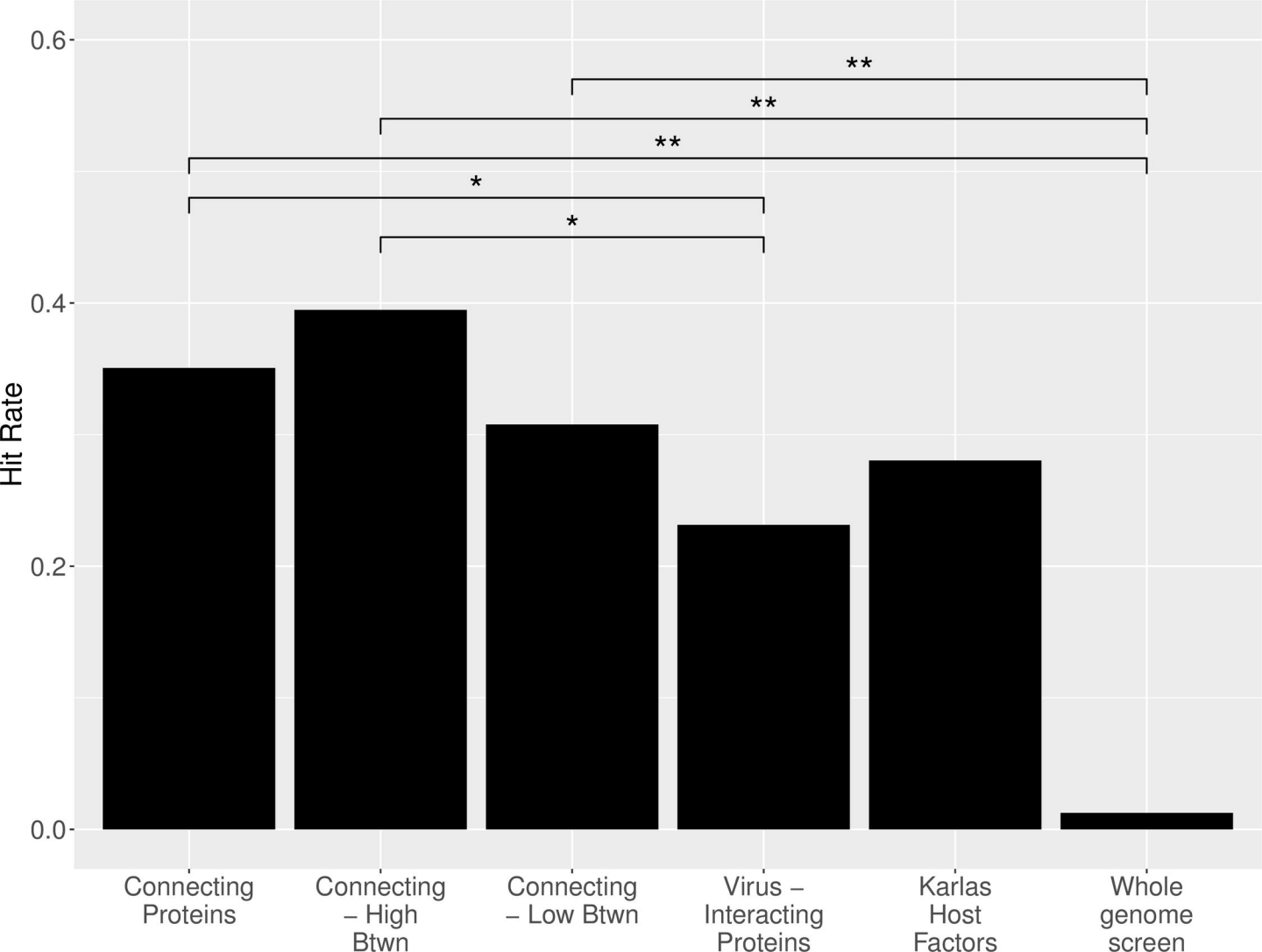

0.6
0.4
0.2
0.0
Hit Rate
**
**
**
*
*
Connecting Proteins
Connecting – High Btwn
Connecting – Low Btwn
Virus – Interacting Proteins
Karlas Host Factors
Whole genome screen

Table 1 Functional Enrichment Analysis of Virus Subnetwork. Functional enrichment analysis of connecting proteins within the virus subnetwork. Proteins were analyzed using DAVID.

Cluster	Number of GO terms	Enrichment Score
Transcription	4	55.4
DNA damage/repair	3	19.2
Protein phosphorylation	19	18.7
Mitosis	5	18.7
Histone reconfiguration	42	14.4
Immune response C-type lectin receptor signaling pathway T cell receptor signaling pathway	3	14.0
Zinc ion binding	4	11.5

Table 2 Functional Enrichment Analysis of Virus Subnetwork. Functional enrichment analysis of virus-interacting proteins within the virus subnetwork. Proteins were analyzed using DAVID.

Cluster	Number of GO terms	Enrichment Score
Ribonucleoprotein/Viral transcription	13	67.4
Cell-cell adhesion	3	46.6
mRNA splicing	9	41.8
Nucleotide binding	10	30.5
Chaperone/UPR	3	22.0
Viral nucleocapsid	3	19.0
mRNA nuclear export	4	17.4
Nucleotide binding/ATP binding	5	17.1
Translation initiation factors	11	13.1
Proteasome/NF-kB MAPK signaling	23	12.0