Contents lists available at ScienceDirect

Seminars in Cell & Developmental Biology

Protein-protein interactions of human viruses

Norman Goodacre [a], Prajwal Devkota [b], Eunhae Bae [a], Stefan Wuchty [b,c,d,e,**], Peter Uetz [f,*]

[a] *Division of Viral Products, Office of Vaccines Research and Review, Center for Biologics Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD, USA*
[b] *Dept. of Computer Science, Univ. of Miami, Coral Gables, FL, 33146, USA*
[c] *Center for Computational Science, Univ. of Miami, Coral Gables, FL, 33146, USA*
[d] *Dept. of Biology, Univ. of Miami, Coral Gables, FL, 33146, USA*
[e] *Sylvester Comprehensive Cancer Center, Miller School of Medicine, University of Miami, Miami, FL, 33136, USA*
[f] *Center for the Study of Biological Complexity, Virginia Commonwealth University, Richmond, VA, 23284, USA*

ARTICLE INFO

ABSTRACT

Viruses infect their human hosts by a series of interactions between viral and host proteins, indicating that detailed knowledge of such virus-host interaction interfaces are critical for our understanding of viral infection mechanisms, disease etiology and the development of new drugs. In this review, we primarily survey human host-virus interaction data that are available from public databases following the standardized PSI-MS format. Notably, available host-virus protein interaction information is strongly biased toward a small number of virus families including herpesviridae, papillomaviridae, orthomyxoviridae and retroviridae. While we explore the reliability and relevance of these protein interactions we also survey the current knowledge about viruses functional and topological targets. Furthermore, we assess emerging frontiers of host-virus protein interaction research, focusing on protein interaction interfaces of hosts that are infected by different viruses and viruses that infect multiple hosts. Finally, we cover the current status of research that investigates the relationships of virus-targeted host proteins to other comorbidities as well as the influence of host-virus protein interactions on human metabolism.

Contents

* Corresponding author.
** Corresponding author at: Dept. of Computer Science, Univ. of Miami, Coral Gables, FL, 33146, USA.
E-mail addresses: wuchtys@cs.miami.edu (S. Wuchty), peter@uetz.us (P. Uetz).

1. Introduction

Bacteria and viruses are the most important pathogens on earth. While most bacteria can be directly eliminated with antibiotics, viruses can only be constrained in their growth, posing a challenge for treatment. This observation is a direct consequence of the fact that viruses are composed of only nucleic acids, proteins, sometimes lipids and a few other compounds. Thus, the survival of viruses is almost entirely dependent on molecular protein-protein interactions (PPI) with their hosts. Furthermore, viral variation occurs rapidly, often with significant adaptation within each host [1]. Hence, strategies for the development of safe antivirals often depend on precise targeting of virus-host PPIs and a deep understanding of viral biology.

In this review we relate the diversity of viruses to their medical importance. We use the number of currently known virus-host PPIs as a proxy for our knowledge of virus biology. Given the vast body of literature about virus-host interactions, we primarily base our review on data of human host-virus interactions available in public databases. Although extensive sequence information from next-generation sequencing studies exists, precise knowledge of virus-host PPIs and thus potential targets for antiviral therapies is rather limited, biased and incomplete. Medically important viruses such as HIV and Influenza have received considerable research attention, while other highly infectious viruses have received relatively little attention, as their spread is geographically limited and focused on a narrow range of hosts. Furthermore, their investigation may prove experimentally difficult or accelerated only recently. For example, the Zika threat is relatively recent, having triggered much research activity only in the last few years.

2. Diversity and morbidity of human viruses

As our interest in viruses is primarily driven by their impact on human health and economic toll we wonder how their medical and economic importance is related to research efforts, reflected by the number of genomes sequenced and the number of host-virus PPIs detected. Unsurprisingly, some viruses received a lot of attention. For example, both HIV and Influenza claim a large number of victims and impose a significant economic burden. In turn, viruses such as Hepatitis, MERS and SARS claim relatively few lives, yet come with a considerable economic price tag (Table 1).

Virus diversity can be measured by sequencing virus isolates from different geographical areas. Such investigations are especially informative for RNA viruses that evolve rapidly, resulting in large sequence diversity. In Table 2, we summarize the 20 most studied virus families as a function of the number of sequenced genomes. Notably, we find more than 7,000 genomes of flaviviridae that include the Zika virus. Furthermore, we count more than 2,000 genomes of retroviridae, including the HIV virus. Sequences of flaviviridae and retroviridae are highly variable, as more than 200,000 sequences in both families cannot be clustered into similar sub-groups. In particular, we applied a similarity threshold of 98% sequence identity with the tool CD-HIT-EST [2], considering genomes with more than 98% sequence identity as a single cluster (Table 2). Cluster representatives are assumed to be the longest sequence in each 98%-similar group. When available, complete genomes from RefSeq viral and neighbor complete genomes [3] are considered cluster representatives.

As another aspect of viral variability, humans are infected by a variety of different viruses. For instance, Wylie et al. found that an average of 5.5 viral genera were found in each of 102 healthy individuals [4]. As only five body habitats were screened, including nose, skin, mouth, vagina, and stool, most people likely carry dozens of different viruses. However, only a few lead to clinical symptoms or disease. Furthermore, Poon et al. [5] found numerous variants of Influenza A in individual human hosts. Such variants showed changing abundance over time and between individuals, reflecting their ability to evolve and adapt quickly to changing hosts and conditions.

3. Virus-human host protein-protein interaction databases

During the last decade numerous protein interactions between human viruses and their host cells have been mapped and more thoroughly investigated. Most of these efforts focused on relatively few viruses, such as Hepatitis C virus [6–9], Human Immunodeficiency Virus [10,11], Influenza A virus [12], herpesviruses [13] such as Epstein-Barr Virus [14,15], as well as Dengue [16] and a few others [17]. As a consequence of many high-throughput screens, databases of PPIs have been filled with tens of thousands

Table 1

Human disease burden by viruses. Infections include infected number of people while morbidity and mortality include those that get sick or die, respectively. Cost is the economic damage of these viral diseases from hospitalization or lost work time. Unless otherwise indicated, figures are yearly.

Virus (species)	Virus (family)	Infections	Morbidity	Mortality	Cost	Refs.
HSV-1/2	Herpesviridae	3.7 B / ~700M	3M/yr US	low	$540 M US	[84–88]
HIV-1/2	Retroviridae	36 M ww	2.1 M/yr ww	25 M ww total[1]	$13.7B US	[89]
Influenza A	Orthomyxoviridae	>30 M US[3]	100-600 K US[4]	50 M 1918[2] ww	$10-90B	[90–92]
Measles morbillivirus	Paramyxoviridae	>20 M ww	250k ww	140-500k[5] ww	$3-7B US	[93]
Hepatitis C	Flaviviridae	60-120 M ww	4M ww	500k ww	$10B[6]	[94,95]
Hepatitis B	Hepadnaviridae	248 M ww/yr, ~2.5 B ww total	350 M ww total	600k ww	$1B US	[96–98]
Zika virus	Flaviviridae	740k SA	>2.6k[7] SA	low	$18B ww	[99]
MERS-CoV	Coronaviridae	2067 ww total	1179 ww total	720 ww total	$15-20B	[100]
SARS-CoV	Coronaviridae	>10k ww total	8098 ww total	774 ww total	$40B ww	[101,102]
Rhinovirus A, B, C	Picornaviridae	1B/year US	10-40% of common colds	low	$20B US	[103]
Norovirus (Norwalk virus)	Caliciviridae	19-21 M US; 685 M ww	699 M ww	570-800 US; 200 K children ww; 219 K ww	$4.2B indirect; $60.3B total ww	[104,105]
Dengue virus	Flaviviridae	390 M ww	96 M ww	10k ww	$2.1B US, SA	[106,107]

[1] Globally, since 1981.
[2] Spanish flu of 1918.
[3] 30 million outpatient visits.
[4] 100-600 thousand hospitalizations.
[5] The death rate is decreasing, from 535,000 deaths in 2000 to 139,300 deaths in 2010.
[6] $10.7 billion in direct medical expenditures in the USA for HCV-related disease from 2010 to 2019.
[7] Cases of microcephaly. K,M,B = thousand, million, billion, ww = worldwide, SA = South America.

Table 2
20 best-studied virus families (by number of genomes sequenced). Sequence numbers as of July, 2016. Clustered sequenced were clustered at ≥98% sequence identity). U/C = un-/ clustered gives the fold-reduction under clustering, indicating the extent of sequence redundancy among complete genomes. Genome data was retrieved from Genbank.

Baltimore Class	Family name	Seqs (unclust.)	Disease examples	U/C	Total complete genomes (unclust.)	Complete genomes (clust.)
III (dsRNA)	Reoviridae	65870	Rare diarrhea	5.50	31945	5803
IV (+ssRNA)	Flaviviridae	225112	Zika	3.88	7837	2019
VII (dsRNA-RT)	Hepadnaviridae*	78558	Hepatitis	3.72	7248	1946
II (ssDNA)	Geminiviridae	13158	—	2.77	6421	2316
IV (+ssRNA)	Picornaviridae	85636	Cold etc	2.30	3447	1500
VI (ssRNA-RT)	Retroviridae	716088	AIDS etc	1.37	2890	2103
II (ssDNA)	Circoviridae	7838	—	4.99	2706	542
V (-ssRNA)	Phenuiviridae	4139	Rift Valley fever	4.37	1678	384
IV (+ssRNA)	Coronaviridae	19164	SARS	4.84	1549	320
IV (+ssRNA)	Potyviridae	16115		1.82	1536	843
I (dsDNA)	Papillomaviridae	17847	Warts, cancer	3.80	1364	359
I (dsDNA)	Polyomaviridae	8604	Rare cancers	7.79	1277	164
V (-ssRNA)	Filoviridae	2165	Ebola	34.03	1259	37
IV (+ssRNA)	Togaviridae	8924	Rubella	9.04	1239	137
V (-ssRNA)	Pneumoviridae	22578	Cold-like	20.18	1231	61
II (ssDNA)	Nanoviridae	3110	—	4.20	1183	282
IV (+ssRNA)	Caliciviridae	32405	gastroenteritis	3.67	1072	292
V (-ssRNA)	Paramyxoviridae	29726	measles	3.08	1008	327
IV (+ssRNA)	Bromoviridae	4677	(plants)	1.99	764	384
V (-ssRNA)	Arenaviridae	2639	e.g. Lassa fever	1.62	758	469

* Hepadnaviruses have an RNA intermediate and thus are not strict DNA viruses.

of virus–human interactions. The IMEx (International Molecular Exchange) consortium of databases has been particularly valuable, as members use a standard format (PSI-MS) for recording metadata for PPIs [18], including experimental details. IMEx members and observers include, notably, BioGRID [19], DIP [20], IntAct [21] and MINT [22] (Table 3). Other generalist PPI databases [23] also collect host-virus interactions but are not further discussed here. Specialized web-based resources have been developed to integrate molecular pathogen-host interactions (PHI) and related data from these PPI databases (generic as well as PHI databases are shown in Table 3).

Some PHI databases specialize on only one specific pathogen species such as HCVpro [24]. A wider range of human specific viruses are covered by VirHostNet [25] and others (Table 3).

The differences between databases often exacerbate the comparison of their data. Especially the lack of experimental details may obscure the nature of an interaction as direct or "indirect", e.g. when a protein is co-purified with other proteins without additional evidence whether it is directly binding to specific members of such a complex. For instance, out of the over 17,000 HIV-1 –

human PPIs reported in HIV-1db [26] as of August, 2017, fewer than 7,000 are physical interactions, and the number of direct (rather than "complexed") interactions is unclear. While clear evidence of direct physical interaction is provided for some PPIs (e.g. "binds", "phosphorylates", "cleaves"), the evidence is weaker for others. Compared to PSI-MS standards, these evidence descriptors were generally more ambiguous. Using a strict set of criteria, the number of physical and direct PPIs between HIV-1 and human proteins is about 1,600. Out of the additional ∼ 5,400 physical interactions, the majority have neither evidence for nor against a direct interaction (e.g. "interacts with", "stabilizes", "recruits"). While NCBI's HIV-1DB is an invaluable resource for HIV-1 researchers, "mining" physical PPIs from the entire database confidently without manual inspection or natural language processing can be difficult.

PHI databases that cover a wide range of viruses and hosts include VirHostNet [25], VirusMentha [27] and HPIDB [28] (Table 3). In addition to data storage, VirusMentha and PHIDIAS [29] offer certain visualization and analysis tools. In addition, PHI databases may contain interactions that are not present in generalist PPI databases as a result of exhaustive mining of

Table 3
Overview of host-pathogen and other protein-protein interaction databases that provide human-virus protein interactions. PHI-PPIs were drawn from databases shown in bold.

database	database type	Human viral species	webpage	Physical PPIs* (March, 2018)	Direct PPIs** (March, 2018)	Ref.
HCVPro	PHI	only HCV	http://www.cbrc.kaust.edu.sa/hcvpro/	618	565	[24]
HIV-1 @NCBI	PHI	only HIV	https://www.ncbi.nlm.nih.gov/genome/viruses/retroviruses/hiv-1/interactions/	6,824	1,594	[26]
PHIDIAS	PHI	37	http://www.phidias.us	***	***	[29]
PHISTO	PHI	****	http://www.phisto.org	****	****	[108]
HPIDB	**PHI**	**43**	http://www.agbase.msstate.edu/hpi/main.html	**19,681**	**10,628**	[28]
VirHostNet	PHI	106	http://virhostnet.prabi.fr	20,674	14,013	[25]
VirusMentha	PHI	98	http://virusmentha.uniroma2.it	10,692	5,863	[27]
DenHunt	PHI	1	http://proline.biochem.iisc.ernet.in/DenHunt/)	1,064	682	[109]
DenvInt	PHI	1	https://denvint.000webhostapp.com	784	784	[65]
BioGRID	**PPI**	**13**	https://thebiogrid.org/	**2,427**	**1,936**	[19]
DIP	**PPI**	**48**	http://dip.mbi.ucla.edu/dip/	**519**	**430**	[20]
IntAct	**PPI**	**142**	https://www.ebi.ac.uk/intact/	**14,282**	**5,934**	[21]
MINT	**PPI**	**67**	https://mint.bio.uniroma2.it/	**6,400**	**2,530**	[22]

* "Physical PPIs" refers to PPIs for which there is experimental evidence of a physical interaction, but absence of evidence for a direct interaction using PSI-MS controlled vocabularies.

** "Direct PPIs" refers to PPIs for which there is experimental evidence of a physical, direct interaction using PSI-MS controlled vocabularies.

*** PPI information and evidence requires manual extraction or text-mining.

**** The PHISTO website was unavailable for the duration of the writing of this review.
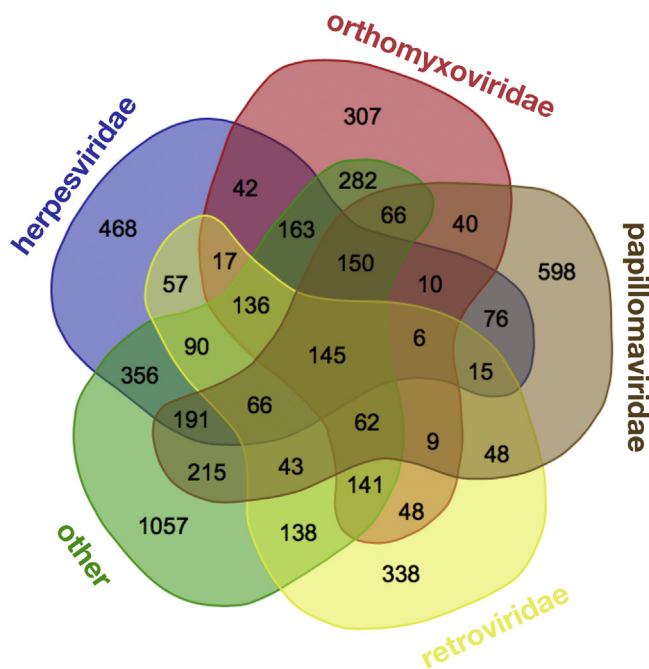
Fig. 1. Human proteins targeted by different virus families. The number of targets is based on all available interactions between viral and human host proteins, including 1,988 targets of Herpesviridae, 1,624 of Orthomyxoviridae, 1,740 of Papillomaviridae, 1,359 of Retroviridae and a pool of 3,301 targets of other virus families. Only a limited number of human proteins is targeted by many different virus families.

Fig. 2. Known human-virus interactions and sequenced viral genomes of virus families. In particular, the number of sequenced genomes of different virus families correlates with the number of known interactions between proteins of the family specific viruses and the human host (Spearman rank correlation coefficient $r_{SR} = 0.41$, $P < 0.05$). Furthermore, we labeled data points that correspond to virus families in Tables 2 and 4.

pathogen-specific data and literature. The most recent update of VirHostNet contains nearly 22,000 virus-human PPIs, of which 14,000 are direct. However, this diversity of source data does present some drawbacks. Many PHI databases lack a standard (i.e. PSI-MS) format (except HPIDB), making inference about the molecular details of interactions such as evidence codes challenging. Furthermore, filtering criteria for collecting virus-host PPIs are often difficult or impossible to ascertain. Finally, the update cycles can be irregular, and often the underlying pipelines for collecting data change from one version to the next. Generally, host-virus protein interaction data in the above PHI databases are integrated mainly from more general PPI databases using automatic integration tools such as PSICQUIC [30] and by manual literature curation. Hence, we excluded PHI databases that do not follow PSI-MS standards from a deeper analysis.

We collected protein interactions for a variety of viral families using the databases in Table 3 (shown in **bold**) and from general databases such as BioGRID and IntAct. Not surprisingly, our understanding of human-virus PPIs is highly biased towards a few well-studied viruses. For instance, only 5 viruses have more than 1,000 physical interactions, such as Influenza A with 3,746, Epstein-Barr Virus (EBV) with 3,163, HIV-1 with 2,540, Herpesvirus 8 with 1,643, and Hepatitis C with 1,082. Notably, Influenza, HIV and Hepatitis viruses were among the most economic burdensome pathogens (Table 1). The human papilloma viruses totaled 4,645 interactions across 29 different species. Evidence for direct interactions of these top viruses varied, but was notably high for herpesvirus 8 (1,623 direct /1,643 physical) and HIV-1 (2,365 direct /2,540 physical). Considering whole virus families (Table 4), we found a total of 5,957 physical interactions with human host proteins involving proteins of orthomyxoviridae. As indicated previously, such sets of interactions are often dominated by a single virus. In addition, we obtained 462 interactions for virus families with fewer than 100 PPIs. In Fig. 1, we summarized the sets of human proteins that were targeted by different virus families and their substantial overlap. While these numbers are roughly similar,
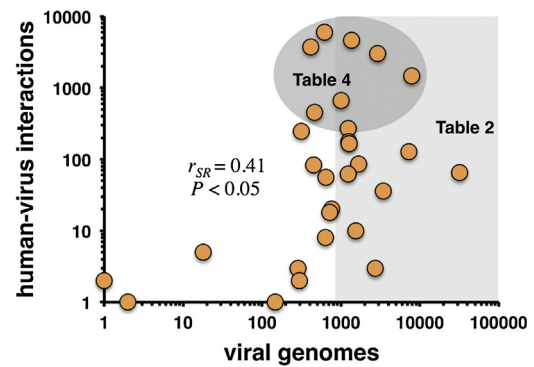
the corresponding virus families represent vastly different genome sizes and virus diversity. For instance, HIV encodes only about 10 proteins with more than 100 interactions per virus protein. EBV, by comparison, encodes about 85 proteins leading to "only" 20 interactions per virus protein on average. On a more quantitative level, the number of interactions between proteins of the human host and viruses that belong to a certain family correlate significantly with the presence of different genomes in a given virus family (Fig. 2). Before we can even begin to interpret these interactions, we need to ask if it is biologically meaningful or even possible if a virus protein has >100 interactions (see below).

4. How reliable are published virus-host interactions?

The large number of PPIs and other high-throughput data raise the question about the reliability of these datasets. We address this problem by examining the reliability of different methods and then ask how many interactions are biologically relevant.

Both yeast two-hybrid (Y2H) and affinity purification coupled to mass spectrometry (AP/MS) approaches have limitations that lead to a significant number of false negatives and false positives [31–33]. While genetic screens using RNA interference (RNAi) or other methods suffer from the same burden (see below) we will discuss these issues in the context of virus-host interactions.

The reliability of PPIs is indicated by the specific evidence for either **direct physical** or **indirect** interaction. For instance, Y2H assays usually yield direct interactions while co-purifications (AP/MS) typically provide indirect interactions, especially when large complexes are studied. Note, that co-purification studies provide direct interactions when only 2 proteins are present in a complex. Around 50% of all virus-human PPIs may be **indirect** (Table 3, bold). While most available virus-host PPIs are physical, it may not be known whether the interaction is direct or indirect. Databases that are based on the PSI-MS format provide experimental methodology information. Of the databases we used, the proportion of interactions that could be inferred as direct ranged from roughly 50% (IntAct, MINT) to around 70% (BioGRID, DIP).

Furthermore, the reliability of the experimental techniques for detecting direct interactions also needs to be considered. Only a few studies exist that systematically validated human-virus interactions for their biochemical or even physiological validity. Among the first attempts to validate human virus-host interactions was our study of interactions between proteins of the human host and the Kaposi sarcoma herpes virus (KSHV) [34]. In particular, we predicted 20 interactions between 8 **KSHV** and 20 human proteins from experimentally determined homologous interactions in yeast,

Table 4

Number of host-virus protein-protein interactions of major human virus families. Interaction numbers are pooled from BioGRID, DIP, HPIDB, IntAct, MINT.

viral family	# virus-human PPIs (physical / direct)	representative virus-human PPIs (physical / direct)
Herpesviridae	5957/3570	Herpesvirus 4 / Epstein-Barr (3,163/1,049); Herpesvirus 8 (1,643/1,623)
Papillomaviridae	4645	Papillomavirus types 1a,3,5,6,6b,8,9,11,16,18,32,33,39 (4,275/2,649)
Orthomyxoviridae	3748/953	Influenza A (3,746/952)
Retroviridae	2998	HIV-1 (2,540/2,365); Primate T-lymphotropic Virus 1 (254/240)
Flaviviridae	1475	Hepatitis C (1,082/802); Dengue (535/535)
Paramyxoviridae	665	measles (481/445); Nipah Henipavirus (133/2)
Adenoviridae	451	Adenovirus types 2,5,12 (378/211)
Pneumoviridae	270	Respiratory Synctial Virus A2 (262/258)
Poxviridae	247	Vaccinia virus (190/47); Variola virus (18/18)
Filoviridae	177	Ebola virus (154/11); Marburg virus (23/0)
Polyomaviridae	165	Macaca Mulatta Polyomavirus 1 (79/65); JC Polyomavirus (41/1); Human Polyomavirus 1 (39/0)
Hepadnaviridae	128	Hepatitis B (127/111)

worm, and fly. Nineteen of these 20 virus-host interactions were tested by CoIP, and an unexpectedly large percentage of interactions (13 out of 19, or 68%) was confirmed.

In another study, Zhang et al. found 109 interactions between 33 **Vaccinia** and ~160 human proteins of which 27 were tested by GST pull-downs [35]. 17 of these were confirmed, translating to a 63% validation rate. While these numbers appear to be rather high, only certain subsets were selected, putatively representing a biased validation rate of a complete Y2H data set.

Khadka et al. [16] screened all 10 proteins of **Dengue virus** against a human liver Y2H library and found 139 interactions involving 8 Dengue and 105 human proteins. 33 interactions were detected by two or more Dengue protein fragments. 16 out of 23 tested interactions were also confirmed by an independent split-luciferase assay. 23 human genes were tested by siRNA assays for their effect on viral replication, and 12 of them were found to be required for Dengue replication. Interestingly, over 40% of the human proteins reported in this study to interact with proteins of the Dengue virus have been implicated in the life cycles of at least one other virus, with the greatest overlap with proteins linked to HCV infection.

In a similar study, 95 human proteins were identified in yeast two-hybrid screens with **Influenza** virus (INFV), and all were tested for their effects on INFV replication in similar siRNA experiments [12]. Of these, three were required for INFV replication and eight exerted a negative effect, for a total of 11 (12%) proteins that affected INFV replication [12]. Similarly, in large-scale siRNA screens for host factors affecting viral replication, the average hit rate was 1% [36,37].

Jäger et al. used AP-MS to identify approximately 2,500 human proteins forming more than 10,000 physical interactions with the ~10 **HIV** proteins analyzed [11]. However, they rigorously filtered their interaction data to obtain a "core" data set of 497 PPIs. Such a set was then compared to literature-curated PPIs in VirusMint [38], that were mostly derived from small-scale targeted studies. Notably, the overlap was only 19 interactions. As another control, Jäger et al. also carried out their purifications in two different cell lines, HEK293 and Jurkat, yet only 196 out of 497 interactions were found in both (39%). Finally, Jäger et al. compared their interaction data with four published RNAi screens [36,39–41] that found a total of 1,071 human genes, of which only 55 overlapped with the 435 proteins found by AP/MS. In a follow-up study, Emig-Agius et al. [42] identified a set of 554 human proteins with "close proximity" to HIV proteins in both AP/MS data and RNAi screens. That is, 382 were direct HIV interactors identified by AP-MS, 79 were identified directly by RNAi, and 148 were novel predictions identified by proximity in the interaction network. These 554 human proteins were also enriched for 40 protein complexes from the CORUM database of human protein complexes [43]. Out of these 40 complexes, 27 had not been reported in previous studies of HIV, while the remaining 13 complexes had been identified by at least one previously

published analysis. Furthermore, 36 of the 40 complexes in the map had at least one subunit that directly interacted with HIV. As a result they found three protein complexes which had not been identified in previous analyses. Additional RNAi experiments confirmed that some subunits were required for efficient HIV infection.

Such examples demonstrate that many human proteins have been identified that are involved in virus infection and physically interact with virus proteins. However, there is little reproducibility even with RNAi screens that have identified human genes required for infection. For instance, a comparison of the first three screens for human proteins involved in HIV infection [36,41,44] revealed only three genes that were identified by all screens.

5. How many interactions does a virus require?

The sheer number of host-virus interactions that have been found for many viruses may suggest that we have identified most if not all interactions. As indicated above, we know from other studies that interaction screens likely contain a large number of false positives. While viral proteins are highly enriched for disordered regions that are known to favor interactions with many partners [45], the question remains as to how many interactions does a virus realistically utilize to infect its host? How do we identify the physiological interactions among those that have been found overall? Unfortunately, such estimates currently do not exist, especially if we assume that viruses of similar proteome size may use similar numbers of PPIs. While it is hard to estimate such numbers for human viruses, bacteriophages may serve as a simpler model: 50 years of research have identified about 30 host-virus interactions between *E. coli* and phage lambda, which encode ~4,000 and 73 proteins, respectively [46]. Assuming that most (physiologically relevant) host-virus interactions in this system have been identified, a large-scale analysis of *E. coli*-lambda interactions revealed 62 interactions in a high-confidence set [47]. However, of the 62 high-confidence PPIs only two were previously known to be physiological, while the role of the other 60 remains unknown. Such observations may indicate that up to 97% of these PPIs are false positives. However, lambda is unusual as compared to other phage in that many lambda proteins are processed during maturation and thus interactions are more difficult to detect. Protein processing seems to be less common among other phages, such as T7, whose 55 proteins are known to be involved in only 15 interactions with its host [46]. For T7 and lambda about 30–40 interactions have been found between their virion proteins, which are easier to detect and possibly more abundant than in human viruses, given the more elaborate virus structure in these tailed phage when compared to the often simple-structured human viruses. Although lambda and T7 may not be the most meaningful references for human viruses they give us a sense for the complexity of the problem. Interestingly, human viruses include many cases with substantial proteolytic processing, such as Hepatitis C Virus, in which a polyprotein prevents

many interactions that are found when the proteolytic fragments are used for interaction screens [48].

6. What are the protein targets of human viruses?

The investigation of different host-virus protein interaction networks revealed functional classes of host proteins that different viruses commonly target. In particular, viruses prefer human host proteins that were involved in cell cycle regulation, signaling, nuclear transport and cell trafficking as well as transcription and translation functions [6,14,34,49,50]. However, different types of viruses will target different human proteins, reminding us that each viral family uses a different strategy to invade a human host cell [51,52].

The abundance of virus-host interaction data prompted topological analysis of networks. For instance, Navratil et al. described a human infectome network (HIN) that linked 416 viral proteins to 1,148 human proteins through 2,099 manually curated virus-host PPIs [53]. In fact, 32% of these cellular proteins are targeted by more than one virus protein of the same virus. A similar fraction, 28% of these cellular targets interact with proteins from more than one virus.

Clearly, virus proteins target a relatively select number of human proteins that are relevant for their replication. These human targets appear to be highly connected: the mean degree of these targets was 38 vs. 10 in non-targeted proteins [53]. Even among highly connected proteins (degree k>5) in the human interactome, the degree of virus targets was twice as large as those of non-targeted proteins. Independent analyses found that viral proteins preferably target human host proteins that are involved in a large number of interactions [14,49,50,54–60]. As the number of interactions is a local measure of centrality, other more global measures of centrality were considered as well. In particular, betweenness centrality measures how many connections go through a particular protein in a network when proteins were mutually connected through their shortest paths. Indeed, various viruses target human host proteins with high betweenness centrality. As a corollary, such central proteins also have significantly shorter paths to other proteins [14,49,50,54–60], as well as participate in a higher number of pathways and protein complexes [61]. Furthermore, topological analyses of revealed that human host proteins that were targeted by viruses are strongly connected among each other [61]. Recently, the focus of modern network research has shifted to the determination of nodes that allow the topological control of a network [62,63]. Such controlling nodes need to be tweaked to exert the behavior of the remaining nodes in the underlying network. Notably, such controlling genes were enriched with essential genes and disease genes, and they appeared in regulatory interactions [64,65]. Furthermore, they also played a role as targeted and required genes of viral infections [66,67]. As a consequence, such centrality measures and the determination of control nodes may allow the computational prediction of potential viral targets based on topological measures [57,58,68–70].

7. Do different virus strains have different interaction patterns?

As discussed above, viruses evolve quickly, and often hundreds if not thousands of strains have been sequenced, suggesting that their diversity may affect their interaction patterns. While several studies have addressed this question, we illustrate the idea with a study of human papilloma viruses (HPV). Neveu et al. [71] tested the interactions of two HPV proteins, E6 and E7, derived from 11 different strains each against 94 and 88 human target proteins, respectively. Each protein often had surprisingly different interaction profiles, depending on the strain from which they were derived. For instance, the affinity of different E6 proteins from different strains for their cellular targets Smad2 and Smad3 differed by more than 10-fold (as measured by an in vitro luciferase protein complementation assay). Interestingly, these authors also showed that the target proteins had very distinct tissue-specific expression profiles, partly explaining the tissue specificity of these viruses. In a complementary study, Gulbahce et al. [72] also investigated the E6 and E7 proteins but focused on the effect these proteins had on the expression levels in different target cells. Both proteins interacted with different proteins but also affected the expression of very different target proteins in human fibroblast and keratinocytes, emphasizing the complexities that arise when virus variation and human cell types are combined.

8. How are protein-protein interactions related to viruses with multiple hosts?

Most viruses are capable of infecting multiple hosts, at least closely related host species. However, some viruses infect quite distantly related species, such as human and bird influenza. One might expect that such "generalist" viruses enjoy an evolutionary advantage due to their potential to infect more hosts. However, in most cases adaptive mutations for one host bring about a decrease in fitness for replicating within another host, with a possible exception when a host's population tends to fluctuate widely [73]. Most viruses require only one host although some viruses (notably arboviruses such as dengue) use humans and other animals as "dead-end" hosts that do not transmit the virus to others in the population [74]. For obligate multi-host viruses, a phenotype specific to each host is often developed, which increases the chances of transmission. The hosts themselves are often very different, in terms of pH, temperature, and cell type. Thus, it is likely that PPIs from obligate multi-host viruses are more diverse than PPIs from optional multi-host viruses.

Arboviruses (arthropod-borne viruses) are among the most burdensome multi-host viruses for humans. As the dengue virus is a good example the DenvInt database [75] catalogues interactions between 10 dengue proteins and both human and mosquito proteins. The dengue-human network consists of 535 interactions between 10 dengue and 335 human proteins, while the dengue-mosquito network consists of 249 interactions between 10 dengue and 140 mosquito proteins.

Even with extensive recombination, as observed among the "chromosomes" of influenza viruses, many viruses are known to jump from animals to humans [76,77], often leading to unusually severe outbreaks. The host environments do not vary as widely for viruses such as influenza A compared to arboviruses. Therefore, host-virus protein interactions PPIs may be more similar. Crossover mutations in influenza viruses occur predominantly in two proteins, hemagglutinin and neuraminidase, which allow for entry and exit, respectively.

Using the virus-human PPIs from the databases in Table 3, the extent of multi-host interactions appears much larger, although their physiological relevance is difficult to determine. In particular, we found 71 human viruses from 15 different viral families in Table 4 (except filoviridae and polyomaviridae) that had interactions with 43 non-human hosts. The most common alternate hosts were simian, avian, mouse, cow and pig. There were over 100 interactions between the 2006 h5n1 pandemic flu strain and chicken. However, there were even some interactions with bacterial and plant hosts (for herpesviral proteins). In all, there were 434 physical interactions, of which 186 were direct. Conversely, non-human viruses were also found to have PPIs with human proteins, with less diversity but overall higher interactions. These 39
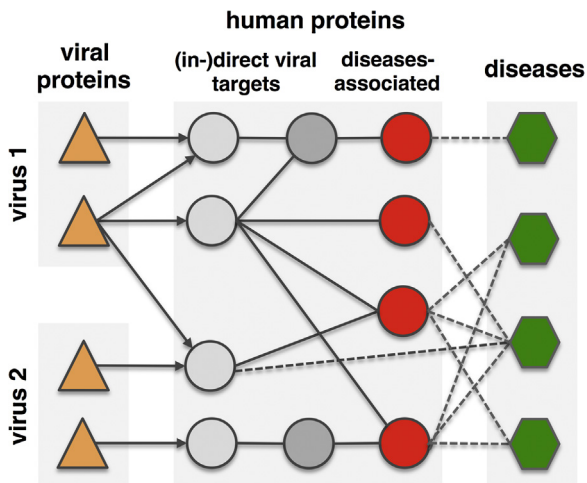
human proteins

Fig. 3. The interactome-diseasome connection. Topological proximity between viral targets and genes associated with virally implicated diseases. Many diseases are directly or indirectly connected to virus proteins and their human targets. Modified after [72].

non-human viruses were from 15 different viral families, of which only 7 were present in Table 4. 627 physical and 525 direct PPIs were found. This greater number of PPIs may be a result of the larger human proteome compared to most alternative hosts. The non-human hosts were similar in both directions of the analysis. However, some differences could be found. Canine viruses were found to have interactions in human, although human viruses were not found to have interactions in canine species. On the other hand, non-human viruses did not come from any bacterial or plant hosts.

9. The virus interactome-diseasome connection

It has been long known that some viruses are involved in diseases not typically associated with infection. For instance, up to 20% of cancers may be caused by viruses such as papilloma or certain herpesviruses [78]. Navratil et al. [53] used a set of virus targets that was compared to a list of 1,729 human genetic disease-related proteins as of the OMIM database [79]. Notably, 13% of human virus targets were also associated with at least one human disease [53], indicating that a human protein interacting with a virus protein is twice as likely to be involved in a disease than a non-target. Most of the diseases found in this study were related to cancer or neurodegenerative diseases. Surprisingly, type-1 diabetes was also associated with virus infection, as were autoimmune diseases in general, as many virus infections elicit a strong immune reaction.

Obtaining a similar conclusion, Gulbahce et al. analyzed the connection between Epstein-Barr-Virus (EBV), human papilloma virus (HPV) and disease [72]. However, these authors not just used protein interactions but also metabolic networks and regulatory interactions (Fig. 3). Using U.S. Medicare patient medical history data derived from 13 million patients, Gulbahce et al. found that many diseases were often associated with infections of the underlying viruses. For instance, patients that suffered from an infection with the human papilloma virus had 15.7 and 2.7 times increased chance of developing retina and bladder cancer and a higher risk of Fanconi anemia [72]. Furthermore, Rozenblatt-Rosen et al. [80] experimentally showed that genomic variations and tumor viruses may cause cancer through related mechanisms. In particular, their systematic analyses of viral host targets identified cancer genes with a success rate that was similar compared to functional genomics and determination of tumor mutations.

10. Virus interactions with the host metabolome

Evidence is mounting that viruses not just highjack host replicative functions, but also the host metabolic machinery. For instance, Adenovirus 5 proteins E4ORF1 and E4ORF6 co-immunoprecipitate with MYC in the nucleus, probably by directly interacting with the cancer protein MYC. While MYC has diverse effects on numerous target genes that it regulates, E4ORF1 induces MYC to activate a subset of glycolytic targets (viruses with a deletion of the E4 protein are defective for inducing glycolysis). Thai et al. [81] conclusively demonstrated that adenovirus induced glycolysis generates metabolites for increased nucleotide biosynthesis in infected cells. Furthermore, Ramière et al. determined a direct interaction between the NS5A of the Hepatitis C virus and cellular hexokinase 2 through protein complementation assays and co-immunoprecipitation. Notably, NS5A expression was sufficient to enhance glucose consumption and lactate secretion. Indicating a direct point of influence on a hosts metabolism through a host-virus protein interaction, viruses putatively manipulate host metabolism to generate more nucleotides and other compounds that are needed for their replication [82]. However, in most host-virus system viral points of metabolic intervention are understudied, and ambiguity remains whether interactions of virus proteins with host enzymes directly or indirectly reprogram metabolism.

11. Conclusions and outlook

Protein-protein interactions are at the core of any virus infection. As a consequence, detailed knowledge of such interactions is critical for our understanding of viral diseases and the development of new drugs. However, knowledge about interactions between host and viral proteins is strongly biased toward a small number of viral families. Notably, these families often have diverse genomes while being of utmost biomedical and economic importance. In turn, for many viruses of lesser medical importance only few interactions are known that are not sufficient to understand their infection mechanisms at this point. Notably, viruses evolve much quicker than their hosts, especially in RNA viruses, not the least because they can produce numerous (variant) virus particles by the time the host can mount an immune response. Thus, viruses can also adapt their host-virus interaction interface faster than a host population can react by mutating its target proteins, although an individual immune system is usually able to fight off an infection [83].

Despite the abundance of known interactions between human and viral proteins, current host-pathogen interaction databases lack the level of specific annotations compared to general protein interaction databases. Furthermore, the level of reliability of host-virus interactions is currently not on par with their intra-species counterparts. As we surmise that host-virus protein interactions are key to our understanding of virus infection and disease, currently known host-virus protein interactions may suggest plausible hypotheses. However, precise mechanisms remain often elusive as the majority of host-virus protein interactions remain to be validated.

While the determination of host-virus protein interaction interfaces is continuing, new frontiers are rapidly emerging. In particular, the availability of host-virus interactions raises the questions of viral co-infections that are reflected by interactions between host and variable viral proteins at the same time. Current knowledge about such co-infections can be inferred from known interactions that have been found from separate analyses that focused on a single host-virus system. However, such interactions will need to be determined in experimental settings that account for the presence of different pathogens. Such settings

will allow the assessment of synergies and interactions between pathogens that putatively have a physiological effect on the host as well as only appear as a consequence of a co-infection. The relationships between diseases that are triggered by host-virus protein interactions and other non-viral comorbidities are increasingly investigated. As a corollary of the previous point, knowledge about such systems is mostly based on assembling and analyzing interaction interfaces that were found independently from the presence of other diseases as well. As a consequence, host-virus interaction interfaces may change in the presence of other diseases. In turn, host-virus interactions may facilitate the emergence of comorbidities, suggesting that host-virus interaction interfaces need to be experimentally assessed in the presence of associated diseases.

While current host-virus interaction interfaces almost entirely focus on interactions between proteins, the impact of viral proteins on enzymes and the human metabolome have hardly been investigated. As a corollary, microbiome studies will identify many more viruses in humans. In the course of these studies, we will also find many more commensal viruses which do interact with their human host but may actually be beneficial, and potentially even help us to fight other pathogens and parasites.

Conflicts of interest

None.

References

[1] K.S. Xue, A.L. Greninger, A. Perez-Osorio, J.D. Bloom, Cooperating H3N2 influenza virus variants are not detectable in primary clinical samples, mSphere 3 (1) (2018).
[2] W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, Bioinformatics 22 (13) (2006) 1658–1659.
[3] J.R. Brister, et al., Virus variation resource–recent updates and future directions, Nucleic Acids Res. 42 (Database issue) (2014) D660–D665.
[4] K.M. Wylie, et al., Metagenomic analysis of double-stranded DNA viruses in healthy adults, BMC Biol 12 (2014) 71.
[5] L.L. Poon, et al., Quantifying influenza virus diversity and transmission in humans, Nat. Genet. 48 (2) (2016) 195–200.
[6] B. de Chassey, et al., Hepatitis C virus infection protein network, Mol. Syst. Biol. 4 (2008) 230.
[7] L.P. Tripathi, et al., Network based analysis of hepatitis C virus core and NS4B protein interactions, Mol. Biosyst. 6 (12) (2010) 2539–2553.
[8] P.T. Dolan, et al., Identification and comparative analysis of hepatitis C virus-host cell protein interactions, Mol. Biosyst. 9 (12) (2013) 3199–3209.
[9] H.T. Ngo, L.V. Pham, J.W. Kim, Y.S. Lim, S.B. Hwang, Modulation of mitogen-activated protein kinase-activated protein kinase 3 by hepatitis C virus core protein, J. Virol. 87 (10) (2013) 5718–5731.
[10] V.W. Gautier, et al., In vitro nuclear interactome of the HIV-1 Tat protein, Retrovirology 6 (2009) 47.
[11] S. Jager, et al., Global landscape of HIV-human protein complexes, Nature 481 (7381) (2011) 365–370.
[12] S.D. Shapira, et al., A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection, Cell 139 (7) (2009) 1255–1267.
[13] E. Fossum, et al., Evolutionarily conserved herpesviral protein interaction networks, PLoS Pathog. 5 (9) (2009) e1000570.
[14] M.A. Calderwood, et al., Epstein-Barr virus and virus human protein interaction maps, Proc. Natl. Acad. Sci. U. S. A. 104 (18) (2007) 7606–7611.
[15] A. Forsman, U. Ruetschi, J. Ekholm, L. Rymo, Identification of intracellular proteins associated with the EBV-encoded nuclear antigen 5 using an efficient TAP procedure and FT-ICR mass spectrometry, J. Proteome Res. 7 (6) (2008) 2309–2319.
[16] S. Khadka, et al., A physical interaction network of dengue virus and human proteins, Mol. Cell. Proteom. 10 (12) (2011), M111 012187.
[17] A. Pichlmair, et al., Viral immune modulators perturb the human molecular network by common and unique strategies, Nature 487 (7408) (2012) 486–490.
[18] S. Orchard, et al., Protein interaction data curation: the International Molecular Exchange (IMEx) consortium, Nat. Methods 9 (4) (2012) 345–350.
[19] A. Chatr-Aryamontri, et al., The BioGRID interaction database: 2013 update, Nucleic Acids Res. 41 (Database issue) (2013) D816–D823.
[20] L. Salwinski, et al., The database of interacting proteins: 2004 update, Nucleic Acids Res. 32 (Database issue) (2004) D449–D451.
[21] S. Orchard, et al., The MIntAct project–IntAct as a common curation platform for 11 molecular interaction databases, Nucleic Acids Res. 42 (Database issue) (2014) D358–D363.
[22] L. Licata, et al., MINT, the molecular interaction database: 2012 update, Nucleic Acids Res. 40 (Database issue) (2012) D857–D861.
[23] D. Szklarczyk, L.J. Jensen, Protein-protein interaction databases, Methods Mol. Biol. 1278 (2015) 39–56.
[24] S.K. Kwofie, U. Schaefer, V.S. Sundararajan, V.B. Bajic, A. Christoffels, HCVpro: hepatitis C virus protein interaction database, Infect. Genet. Evol. 11 (8) (2011) 1971–1977.
[25] T. Guirimand, S. Delmotte, V. Navratil, VirHostNet 2.0: surfing on the web of virus/host molecular interactions data, Nucleic Acids Res. 43 (Database issue) (2015) D583–D587.
[26] D. Ako-Adjei, et al., HIV-1, human interaction database: current status and new features, Nucleic Acids Res. 43 (Database issue) (2015) D566–D570.
[27] A. Calderone, L. Licata, G. Cesareni, VirusMentha: a new resource for virus-host protein interactions, Nucleic Acids Res. 43 (Database issue) (2015) D588–D592.
[28] R. Kumar, B. Nanduri, HPIDB–a unified resource for host-pathogen interactions, BMC Bioinform. 11 (Suppl. 6) (2010) S16.
[29] Z. Xiang, Y. Tian, Y. He, PHIDIAS: a pathogen-host interaction data integration and analysis system, Genome Biol. 8 (7) (2007) R150.
[30] B. Aranda, et al., PSICQUIC and PSISCORE: accessing and scoring molecular interactions, Nat. Methods 8 (7) (2011) 528–529.
[31] Y.C. Chen, S.V. Rajagopala, T. Stellberger, P. Uetz, Exhaustive benchmarking of the yeast two-hybrid system, Nat. Methods 7 (9) (2010) 667–668.
[32] B. Titz, M. Schlesner, P. Uetz, What do we learn from high-throughput protein interaction data? Expert Rev. Proteom. 1 (1) (2004) 111–121.
[33] K. Venkatesan, et al., An empirical framework for binary interactome mapping, Nat. Methods 6 (1) (2009) 83–90.
[34] P. Uetz, et al., Herpesviral protein networks and their interaction with the human proteome, Science 311 (5758) (2006) 239–242.
[35] L. Zhang, et al., Analysis of vaccinia virus-host protein-protein interactions: validations of yeast two-hybrid screenings, J. Proteome Res. 8 (9) (2009) 4311–4318.
[36] A.L. Brass, et al., Identification of host proteins required for HIV infection through a functional genomic screen, Science 319 (5865) (2008) 921–926.
[37] A.L. Brass, et al., The IFITM proteins mediate cellular resistance to influenza A H1N1 virus, West Nile virus, and dengue virus, Cell 139 (7) (2009) 1243–1254.
[38] A. Chatr-aryamontri, et al., VirusMINT: a viral protein interaction database, Nucleic Acids Res. 37 (Database issue) (2009) D669–D673.
[39] R. Konig, et al., Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication, Cell 135 (1) (2008) 49–60.
[40] M.L. Yeung, L. Houzet, V.S. Yedavalli, K.T. Jeang, A genome-wide short hairpin RNA screening of jurkat T-cells for human proteins contributing to productive HIV-1 replication, J. Biol. Chem. 284 (29) (2009) 19463–19473.
[41] H. Zhou, et al., Genome-scale RNAi screen for host factors required for HIV replication, Cell. Host Microbe 4 (5) (2008) 495–504.
[42] D. Emig-Agius, et al., An integrated map of HIV-human protein complexes that facilitate viral infection, PLoS One 9 (5) (2014) e96687.
[43] A. Ruepp, et al., CORUM: the comprehensive resource of mammalian protein complexes–2009, Nucleic Acids Res. 38 (Database issue) (2010) D497–D501.
[44] R. Konig, et al., Human host factors required for influenza virus replication, Nature 463 (7282) (2010) 813–817.
[45] N. Tokuriki, C.J. Oldfield, V.N. Uversky, I.N. Berezovsky, D.S. Tawfik, Do viral proteins possess unique biophysical features? Trends Biochem. Sci. 34 (2) (2009) 53–59.
[46] R. Hauser, et al., Bacteriophage protein-protein interactions, Adv. Virus Res. 83 (2012) 219–298.
[47] S. Blasche, S. Wuchty, S.V. Rajagopala, P. Uetz, The protein interaction network of bacteriophage lambda with its host, Escherichia coli, J. Virol. 87 (23) (2013) 12745–12755.
[48] M. Flajolet, et al., A genomic approach of the hepatitis C virus generates a protein interaction map, Gene 242 (1–2) (2000) 369–379.
[49] M.D. Dyer, T.M. Murali, B.W. Sobral, The landscape of human proteins interacting with viruses and other pathogens, PLoS Pathog. 4 (2) (2008) e32.
[50] S. Wuchty, G. Siwo, M.T. Ferdig, Viral organization of human proteins, PLoS One 5 (8) (2010) e11796.
[51] A.F. Brito, J.W. Pinney, Protein-Protein interactions in virus-host systems, Front. Microbiol. 8 (2017) 1557.
[52] A. Dix, S. Vlaic, R. Guthke, J. Linde, Use of systems biology to decipher host-pathogen interaction networks and predict biomarkers, Clin. Microbiol. Infect. 22 (7) (2016) 600–606.
[53] V. Navratil, B. de Chassey, C.R. Combe, V. Lotteau, When the human viral infectome and diseasome networks collide: towards a systems biology platform for the aetiology of human diseases, BMC Syst. Biol. 5 (2011) 13.
[54] S.D. Durmus Tekir, K. Uelgen, Systems biology of pathogen-host interaction: networks of protein-protein interaction within pathogens and pathogen-human interactions in the post-genomic era, Biotechn. J. 8 (1) (2013) 85–96.
[55] R. Arnold, K. Boonen, M.G. Sun, P.M. Kim, Computational analysis of interactomes: current and future perspectives for bioinformatics approaches to model the host-pathogen interaction space, Methods 57 (4) (2012) 508–518.

[56] M.J. Korth, N. Tchitchek, A.G. Benecke, M.G. Katze, Systems approaches to influenza-virus host interactions and the pathogenesis of highly virulent and pandemic viruses, Semin. Immunol. 25 (3) (2013) 228–239.

[57] E. Nourani, F. Khunjush, S. Durmus, Computational approaches for prediction of pathogen-host protein-protein interactions, Front. Microbiol. 6 (2015) 94.

[58] M.D. Dyer, T.M. Murali, B.W. Sobral, Supervised learning and prediction of physical interactions between human and HIV proteins, Infect. Genet. Evol. 11 (5) (2011) 917–923.

[59] S. Durmus Tekir, T. Cakir, K.O. Ulgen, Infection strategies of bacterial and viral pathogens through pathogen-human protein-protein interactions, Front. Microbiol. 3 (2012) 46.

[60] R.R. Halehalli, H.A. Nagarajaram, Molecular principles of human virus protein-protein interactions, Bioinformatics 31 (7) (2015) 1025–1033.

[61] R. Mariano, S. Khuri, P. Uetz, S. Wuchty, Local action with global impact: highly similar infection patterns of human viruses and bacteriophages, mSystems 1 (2) (2016).

[62] M. Ishitsuka, T. Akutsu, J.C. Nacher, Critical controllability in proteome-wide protein interaction network integrating transcriptome, Sci. Rep. 6 (2016) 23541.

[63] J.C. Nacher, T. Akutsu, Analysis of critical and redundant nodes in controlling directed and undirected complex networks using dominating sets, J. Compl. Networks 2 (4) (2014) 394–412.

[64] S. Wuchty, Controllability in protein interaction networks, Proc. Natl. Acad. Sci. U. S. A. 111 (19) (2014) 7156–7160.

[65] S. Khuri, S. Wuchty, Essentiality and centrality in protein interaction networks revisited, BMC Bioinform. 16 (2015) 109.

[66] S. Wuchty, T. Boltz, H. Kucuk-McGinty, Links between critical proteins drive the controllability of protein interaction networks, Proteomics 17 (10) (2017) 1–11.

[67] A. Vinayagam, et al., Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets, Proc. Natl. Acad. Sci. U. S. A. 113 (18) (2016) 4976–4981.

[68] R. Mariano, S. Wuchty, Structure-based prediction of host-pathogen protein interactions, Curr. Opin. Struct. Biol. 44 (2017) 119–124.

[69] T.M. Murali, M.D. Dyer, D. Badger, B.M. Tyler, M.G. Katze, Network-based prediction and analysis of HIV dependency factors, PLoS Comput. Biol. 7 (9) (2011) e1002164.

[70] O. Tastan, Y. Qi, J.G. Carbonell, J. Klein-Seetharaman, Prediction of interactions between HIV-1 and human proteins by information integration, Pac. Symp. Biocomput. (2009) 516–527.

[71] G. Neveu, et al., Comparative analysis of virus-host interactomes with a mammalian high-throughput protein complementation assay based on Gaussia princeps luciferase, Methods 58 (4) (2012) 349–359.

[72] N. Gulbahce, et al., Viral perturbations of host networks reflect disease etiology, PLoS Comput. Biol. 8 (6) (2012) e1002531.

[73] S.F. Elena, P. Agudelo-Romero, J. Lalic, The evolution of viruses in multi-host fitness landscapes, Open Virol. J. 3 (2009) 1–6.

[74] S.C. Weaver, A.D. Barrett, Transmission cycles, host range, evolution and emergence of arboviral disease, Nat. Rev. Microbiol. 2 (10) (2004) 789–801.

[75] L. Dey, A. Mukhopadhyay, DenvInt: a database of protein-protein interactions between dengue virus and its hosts, PLoS Negl. Trop. Dis. 11 (10) (2017) e0005879.

[76] J.N. Mandl, et al., Reservoir host immune responses to emerging zoonotic viruses, Cell 160 (1–2) (2015) 20–35.

[77] D. Vijaykrishna, R. Mukerji, G.J. Smith, RNA virus reassortment: an evolutionary mechanism for host jumps and immune evasion, PLoS Pathog. 11 (7) (2015) e1004902.

[78] A. Morales-Sanchez, E.M. Fuentes-Panana, Human viruses and cancer, Viruses 6 (10) (2014) 4047–4079.

[79] V.A. McKusick, Mendelian inheritance in man and its online version, OMIM, Am. J. Hum. Genet. 80 (4) (2007) 588–604.

[80] O. Rozenblatt-Rosen, et al., Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins, Nature 487 (7408) (2012) 491–495.

[81] M. Thai, et al., Adenovirus E4ORF1-induced MYC activation promotes host cell anabolic glucose metabolism and virus replication, Cell. Metab. 19 (4) (2014) 694–701.

[82] S.J. Miyake-Stoner, C.C. O'Shea, Metabolism goes viral, Cell. Metab. 19 (4) (2014) 549–550.

[83] A. Christiaansen, S.M. Varga, J.V. Spencer, Viral manipulation of the host immune response, Curr. Opin. Immunol. 36 (2015) 54–60.

[84] S.B. Boppana, K.B. Fowler, Persistence in the population: epidemiology and transmisson, in: A. Arvin, G. Campadelli-Fiume, E. Mocarski, P.S. Moore, B. Roizman, R. Whitley, K. Yamanishi (Eds.), Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis, Cambridge, 2007.

[85] S. Burrel, et al., Ancient recombination events between human Herpes simplex viruses, Mol. Biol. Evol. 34 (7) (2017) 1713–1721.

[86] C. Johnston, S.L. Gottlieb, A. Wald, Status of vaccine research and development of vaccines for herpes simplex virus, Vaccine 34 (26) (2016) 2948–2952.

[87] K.J. Looker, et al., Global and regional estimates of prevalent and incident Herpes simplex virus type 1 infections in 2012, PLoS One 10 (10) (2015) e0140765.

[88] T.D. Szucs, K. Berger, D.N. Fisman, S. Harbarth, The estimated economic burden of genital herpes in the United States. An analysis using two costing approaches, BMC Infect. Dis. 1 (2001) 5.

[89] N.A. Menzies, et al., The cost of providing comprehensive HIV treatment in PEPFAR-supported programs, AIDS 25 (14) (2011) 1753–1760.

[90] N.P. Johnson, J. Mueller, Updating the accounts: global mortality of the 1918-1920 Spanish influenza pandemic, Bull. Hist. Med. 76 (1) (2002) 105–115.

[91] M.H. Merson, J. O'Malley, D. Serwadda, C. Apisuk, The history and challenge of HIV prevention, Lancet 372 (9637) (2008) 475–488.

[92] N.A. Molinari, et al., The annual impact of seasonal influenza in the US: measuring disease burden and costs, Vaccine 25 (27) (2007) 5086–5096.

[93] M.K. Patel, et al., Progress toward regional measles elimination - worldwide, 2000-2015, MMWR Morb Mortal Wkly. Rep. 65 (44) (2016) 1228–1233.

[94] J.B. Wong, G.M. McQuillan, J.G. McHutchison, T. Poynard, Estimating future hepatitis C morbidity, mortality, and costs in the United States, Am. J. Public Health 90 (10) (2000) 1562–1569.

[95] H. Wang, Collaborators GBoD, Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015, Lancet 388 (10053) (2016) 1459–1544.

[96] J.E. Maynard, Hepatitis B: global importance and need for control, Vaccine 8 (1990), Suppl:S18-20; discussion S21-13.

[97] J.J. Ott, G.A. Stevens, J. Groeger, S.T. Wiersma, Global epidemiology of hepatitis B virus infection: new estimates of age-specific HBsAg seroprevalence and endemicity, Vaccine 30 (12) (2012) 2212–2219.

[98] K. Keshavarz, et al., Economic burden of hepatitis B virus-related diseases: evidence from iran, Hepat. Mon. 15 (4) (2015) e25854.

[99] M. Fischer, Zika Virus Epidemiology Update, 2016.

[100] X.C. Tang, et al., Identification of human neutralizing antibodies against MERS-CoV and their role in virus adaptive evolution, Proc. Natl. Acad. Sci. U. S. A. 111 (19) (2014) E2018–E2026.

[101] K. Oberholtzer, Learning from SARS: Preparing for the Next Disease Outbreak–Workshop Summary, National Academies Press, 2004.

[102] J.-W. Lee, W.J. McKibbin, Estimating the global economic costs of SARS, in: Learning from SARS: Preparing for the Next Disease Outbreak: Workshop Summary, Institute of Medicine (US) Forum on Microbial Threats, National Academies Press, Washington, DC, 2004.

[103] M. Simasek, D.A. Blandino, Treatment of the common cold, Am. Fam. Phys. 75 (4) (2007) 515–520.

[104] S.M. Bartsch, B.A. Lopman, S. Ozawa, A.J. Hall, B.Y. Lee, Global economic burden of norovirus gastroenteritis, PLoS One 11 (4) (2016) e0151219.

[105] E. Robilotti, S. Deresinski, B.A. Pinsky, Norovirus, Clin. Microbiol. Rev. 28 (1) (2015) 134–164.

[106] D.S. Shepard, L. Coudeville, Y.A. Halasa, B. Zambrano, G.H. Dayan, Economic impact of dengue illness in the Americas, Am. J. Trop. Med. Hyg. 84 (2) (2011) 200–207.

[107] S. Bhatt, et al., The global distribution and burden of dengue, Nature 496 (7446) (2013) 504–507.

[108] S. Durmus Tekir, et al., PHISTO: pathogen-host interaction search tool, Bioinformatics 29 (10) (2013) 1357–1358.

[109] P. Karyala, et al., DenHunt - a comprehensive database of the intricate network of dengue-human interactions, PLoS Negl. Trop. Dis. 10 (9) (2016) e0004965.