Title

Systems Biology of Virus-Host Protein Interactions: From Hypothesis Generation to Mechanisms of Replication and Pathogenesis

Journal

Annual Review of Virology, 9(1)

ISSN

2327-056X

Authors

Shah, Priya S
Beesabathuni, Nitin S
Fishburn, Adam T
et al.

Publication Date

2022-09-29

DOI

10.1146/annurev-virology-100520-011851

Peer reviewed

HHS Public Access

Author manuscript

Annu Rev Virol. Author manuscript; available in PMC 2023 May 01.

Published in final edited form as:

Annu Rev Virol. 2022 September 29; 9(1): 397–415. doi:10.1146/annurev-virology-100520-011851.

Systems Biology of Virus-Host Protein Interactions: From Hypothesis Generation to Mechanisms of Replication and Pathogenesis

Priya S. Shah[1,2], **Nitin S. Beesabathuni**[2,*], **Adam T. Fishburn**[1,*], **Matthew W. Kenaston**[1,*], **Shiaki A. Minami**[2,*], **Oanh H. Pham**[1,*], **Inglis Tucker**[1,*]

[1]Department of Microbiology and Molecular Genetics, University of California, Davis, California, USA

[2]Department of Chemical Engineering, University of California, Davis, California, USA

Abstract

As obligate intracellular parasites, all viruses must co-opt cellular machinery to facilitate their own replication. Viruses often co-opt these cellular pathways and processes through physical interactions between viral and host proteins. In addition to facilitating fundamental aspects of virus replication cycles, these virus-host protein interactions can also disrupt physiological functions of host proteins, causing disease that can be advantageous to the virus or simply a coincidence. Consequently, unraveling virus-host protein interactions can serve as a window into molecular mechanisms of virus replication and pathogenesis. Identifying virus-host protein interactions using unbiased systems biology approaches provides an avenue for hypothesis generation. This review highlights common systems biology approaches for identification of virus-host protein interactions and the mechanistic insights revealed by these methods. We also review conceptual innovations using comparative and integrative systems biology that can leverage global virus-host protein interaction data sets to more rapidly move from hypothesis generation to mechanism.

Keywords

systems biology; proteomics; protein interactions

INTRODUCTION

Viruses are basic parasites of host machinery. Even the largest DNA viruses rely on host translation and energy metabolism machinery. Smaller viruses, which must be more economical with their genomic space, make use of even more host machinery at multiple steps in the general virus replication cycle. Thus, viruses can co-opt host proteins for entry (surface receptors, trafficking factors), genome replication and translation (polymerases,

prsshah@ucdavis.edu .

*These authors contributed equally to this work

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

translation factors), and egress (assembly, trafficking factors). Because of this fundamental molecular reliance of viruses on their hosts, all viruses interact with their hosts in the most intimate of ways, through virus-host protein interactions.

Major advances in molecular biology, biochemistry, cell biology, and analytical methods have enabled the systematic and comprehensive discovery of virus-host protein interactions at an unprecedented scale. Broadly speaking, our ability in the field to identify such virus-host protein interactions far exceeds our capacity to understand their consequences. Many of the tools to transform protein interaction data sets into mechanistic insights exist. However, integrating them requires coordinated design and execution.

In this review, we discuss the state of the field of systems biology of virus-host protein interactions. We first discuss how different systems biology technologies are used to identify virus-host protein interactions, including the advantages and disadvantages of these approaches, and newer technologies that could have large impacts on this field. We then discuss conceptual innovations that have been applied to the systems-level study of virus-host protein interactions. This includes comparative and integrative approaches that can help the field move from systems to function and disease more rapidly. We focus on key studies spanning several different viruses to illustrate these specific points.

TOOLS

Several high-throughput techniques exist for identifying virus-host protein interactions. In this section, we review the fundamentals of each technique, the advantages and disadvantages of the techniques, and new technologies that could further transform the identification of virus-host protein interactions. We also highlight virus-host protein interaction studies that use these approaches, with a focus on seminal studies in the field.

Yeast Two-Hybrid

Yeast two-hybrid (Y2H) screening is a more than 30-year-old technology that revolutionized the study of protein interactions by creating a genetics-based functional readout of this phenomenon in yeast, a high-throughput model organism (1). The premise of this technology relies on separating two functional domains of the Gal4 transcription activator (DNA binding and transcription activation) and fusing them to complementary DNA (cDNA) libraries of bait protein and prey, respectively. A successful bait-prey protein interaction will drive the expression of a reporter gene by bringing together the Gal4 functional domains at the upstream activating sequence (UAS) promoter. The reporter gene can be carefully chosen to confer survival in minimal media so that a direct protein interaction can be read out through selection strategies (Figure 1a). While Y2H was originally developed for soluble protein domains through nuclear activity, modifications have enabled membrane Y2H (2, 3) and a pooled high-throughput sequencing readout (4, 5).

Y2H allows for identification of direct interactions between bait and prey in a genetically tractable model organism, which has advantages and disadvantages. As a high-throughput technology, it is valuable for comprehensive efforts to map virus-host protein interactions for large DNA viruses, such as Epstein-Barr virus (6). Because Y2H relies on heterologous

expression of cDNA libraries in yeast, it opens up the study of many virus-host protein interactions, even if tractable cell culture models and/or high-quality proteomes are not readily available, as is the case for many viruses that infect arthropods (7, 8). Identification of direct one-to-one interactions is biochemically advantageous because it is more likely that the interaction can be dissected in vitro if the interaction is direct. However, this also means that the existence of protein complexes, or disruption of these complexes, remains out of reach with Y2H technology.

Classic Affinity Purification and Mass Spectrometry

Mass spectrometry (MS) can be used for global identification of hundreds or even thousands of proteins from a single sample. Protein samples are digested into smaller peptide fragments by well-defined proteases. These peptide fragments are then ionized and analyzed for their unique spectra of mass-to-charge ratio. These spectra are then reverse searched against a theoretical database of spectra to identify proteins in the sample (9). When paired with affinity purification (AP) of a target protein of interest (bait), this approach enables discovery of protein interactions (prey), including complexes and interactions that may not be direct (Figure 1b). Efficient purification of the bait is essential for AP-MS because proteins cannot be amplified in vitro. High-quality antibodies are useful but not always available. Moreover, using a bait-specific antibody creates challenges when comparing protein interactions across several baits, due to differences in antibody performance. Consequently, affinity tag approaches have gained popularity. These affinity tag approaches can be used in the context of infection if an infectious clone can be made (10–13) or with ectopic expression of individual viral proteins as bait.

Generating infectious viral clones with affinity tags can reveal interactions that occur during the context of infection. One elegant application of this approach by Ileana Cristea and colleagues (14) revealed novel mechanisms by which human cytomegalovirus (HCMV), a large DNA virus, regulates its complex gene expression program through virus-host protein interactions. The authors used an HCMV infectious pUL83 mutant with two different affinity tags to identify pUL83-host protein interactions during infection. Ultimately, host protein IFI16 was shown to regulate immediate early gene expression through a protein interaction with pUL83 by recruiting IFI16 to immediate early promoters. Thus, the natural antiviral gene expression program is repurposed by HCMV to promote virus replication (15). It should be noted that inserting affinity tags into infectious clones can be challenging, especially for small RNA viruses that often eliminate these insertions during error-prone replication. While affinity-tagged infectious clones have been generated for some RNA viruses (10, 12, 13, 15A, 19, 79), these methods are more commonly used for large DNA viruses that can tolerate these insertions more readily (14, 15B, 15C, 15D, 15E, 33).

In contrast to AP-MS approaches in the context of infection, ectopic expression approaches have the advantage of rapidly generating protein interaction data across many different conditions or viral proteins. Three comprehensive studies of virus-host protein interaction mapping using ectopic expression and AP-MS, all published in 2012, revealed mechanistic insight into how viruses hijack host processes (16–18). Together, they mark the beginning of this era. In a study on human immunodeficiency virus (HIV)-host protein interactions,

Stephanie Jäger and colleagues (17) produced the first comprehensive virus-host protein interaction network for a single virus. In a partnering publication and concurrent with a complementary study focused on Vif-host protein interactions identified during HIV infection, the proteomic data were leveraged to identify CBFβ as a critical complex member to reconstitute the Vif-APOBEC3G complex in vitro (19, 20). In another study focused on the immune-modulating functions of viral proteins, Andreas Pichlmair and colleagues (18) identified immune-related virus-host interactions across a diverse set of DNA and RNA virus families. An antiviral activity was identified for the host protein hnRNPU. Additionally, the WNK kinase family was found to interact with vaccinia virus K7 protein, which inhibited the antiviral activity of WNK proteins. Finally, Orit Rosenblatt-Rosen and colleagues (16) systematically identified virus-host protein interactions for oncogenic viruses to reveal common pathways targeted by these viruses and improve cancer gene identification.

Proximity Labeling Proteomics

While traditional AP-MS approaches have been used with great success to uncover virus-host protein interactions, several recent advances in in situ proximity labeling can improve the search for one-to-many virus-host protein interactions. Traditional AP-MS techniques must be thorough enough to capture sufficient bait while being biochemically gentle to prevent breakdown of bait-prey interactions. This can pose a challenge for detecting transient interactions, or those that occur in biochemically challenging cellular compartments, such as within membranes. The development of proximity labeling–based proteomics circumvents these challenges. Rather than identify prey proteins via purification of multi-protein complexes in vitro, complexes are labeled in situ in live cells or organisms and prey are purified by AP of the label. Labeling proteins in situ has the added advantage of avoiding biologically irrelevant interactions that may occur when subcellular compartments are mixed after lysis in an in vitro approach.

Proximity-based proteomic approaches are broadly split into two main methods, both of which rely on biotin labeling and the high-affinity interaction between biotin and streptavidin for purification (Figure 1c). In the first, a promiscuous mutant of the *Escherichia coli* biotin ligase BirA was engineered to directly biotinylate substrates (21, 22). The major drawback of initial iterations of first-generation BioID systems was a long labeling time (>12 h), which prevented the capture of dynamic processes such as the early stages of viral infection. Pioneering work by Alice Ting's group (23) has led to a faster labeling method, using an engineered ascorbate peroxidase enzyme (APEX). In the presence of hydrogen peroxide, APEX quickly converts biotin-phenol into a biotin-phenoxyl radical that spontaneously associates with tyrosine residues before long-scale diffusion can occur, allowing labeling within an ~20-nm radius within minutes, rather than hours (21). Since their initial introduction, both BioID and APEX approaches have been continuously refined by Alice Ting and Kyle Roux to reduce labeling time (BioID2, TurboID, APEX2) and decrease tag size (miniTurbo), which can introduce sterics that affect bait function and interaction with prey (23–25). An increasing number of virus-host protein interaction studies are being published using proximity labeling as this technology matures (26–34).

Scoring Protein Interaction Data

The sensitivity of MS-based proteomics inevitably creates a trade-off between signal and noise. Fortunately, thoughtful experimental design and rigorous proteomic scoring can overcome many of these issues.

General approaches.—Common proteomic background can be identified in the Contaminant Repository for Affinity Purification (CRAPome) database (35). This database compiles proteomic data for many common protein purification systems, including affinity tags (e.g., FLAG, Strep) and proximity labeling (e.g., BioID). Simple thresholds can be assigned to remove promiscuous background.

Systematic approaches.—For ectopic expression, systematic comparison of prey across many viral baits can reveal highly specific protein interactions. Proteomic scoring algorithms that emphasize specificity of the prey across different baits, such as Mass spectrometry interaction STatistics (MiST), capture these highly specific protein interactions. In the development of MiST, specificity of the interaction was empirically shown to be a major predictor of biologically relevant virus-host interactions by comparing the predicted high-scoring interactions to gold standard interactions previously validated for HIV (17). Intuitively, this emphasis on specificity also makes sense for small RNA viruses such as HIV. Their genetic economy means that it is unlikely that multiple proteins will evolve overlapping functions. Thus, we expect the true protein interaction landscape of each viral bait to be fairly unique. In contrast, large DNA viruses employing MiST scoring have a decreased emphasis on specificity (36). Their larger genome size means that some functions may overlap.

Quantitative approaches.—Proteomic studies focused on protein interactions for a single viral protein or host factor during infection result in a more limited data set. While specificity cannot be leveraged in the same way as in systematic approaches, quantitative data acquisition and specialized scoring systems can still help remove proteomic background and identify proteins of interest. Significance Analysis of INTeraction (SAINT) (37) uses label-free quantification of proteins (spectral counts with SAINT or fragment intensity with SAINTq) (38) to assign each interaction to a probability distribution that is used to estimate the likelihood of a true interaction. SAINT is especially beneficial for smaller data sets because it uses all available data for each bait-prey pair to infer likelihood of true and false interactions. Protein interactions that are significantly altered over the course of an infection can also be identified through quantification and statistical analysis of peptide/fragment abundance under different conditions. Several groups have started using SAINT to score virus-host AP-MS data (34, 39–42) and quantitative targeted proteomics to identify significant changes in these virus-host protein interactions (43, 44).

Opportunities for Discoveries on a Larger Scale

Defining the comprehensive landscape of protein interactions for a virus requires systematically repeating this process for each viral protein. For viruses with small genomes, this can be done realistically through ectopic expression or the creation of infectious clones if the virus is genetically tractable. While heroic efforts have resulted in comprehensive

virus-host protein interaction networks for large DNA viruses using systematic AP-MS approaches (36, 45), the enormous scale of these projects limits their number. Consequently, scalable methods to identify protein interactions would be incredibly valuable.

One recent advancement in MS-based identification of protein interactions without the need for AP uses thermal proximity coaggregation (TPCA) (46). A melt curve can be generated for protein complexes by using MS to identify proteins present in soluble and aggregated fractions across a temperature gradient. Proteins that interact have similar melt curves, and protein complexes can be inferred in this manner. A major advantage of TPCA is that protein interactions can be resolved independent of bait protein purification, thus enabling proteome-wide analysis of protein complexes in a single experiment. It also can resolve how multiple proteins interact together as a complex, which would require double purification or other biochemical approaches to resolve by traditional AP-MS. However, the interactions are inferred indirectly and must be validated to some degree. A recent study by Joshua Justice and colleagues (47) involved TPCA analysis on herpes simplex virus (HSV)-infected cells. The authors were able to uncover previously validated and novel host-host and virus-host protein interactions. They resolved the temporal dynamics of multi-protein complexes, including those involved in viral DNA sensing. TPCA promises to be a rich area in the study of virus-host protein interactions.

Protein-protein interaction sequencing (PROPER-seq) is another method that relies on barcoding and high-throughput sequencing to recover protein-protein interactions en masse (48). Briefly, PROPER-seq involves the conversion of the transcriptome of the host cells into libraries of protein-messenger RNA barcode fusions. The libraries are then reverse transcribed and mixed in vitro, and the DNA barcodes of interacting proteins are ligated and recovered by high-throughput sequencing. This approach can be extended for profiling virus-host protein interactions by generating barcoded libraries from infected cells and could be a powerful tool. However, characterization of protein interactions in vitro [e.g., without post-translational modifications (PTMs)] is a major limitation of this approach that should be considered carefully.

COMPARATIVE APPROACHES

Here we review how identifying virus-host protein interactions in multiple systems biology studies can be leveraged to improve data quality, identify how host networks are hijacked, and provide evolutionary insights.

Meta-analysis

As more comprehensive data sets become available, it is valuable to compare these studies to determine the extent to which interactions depend on screening technique, cell line, virus strain, or scoring approach. Two previously published studies on Zika virus (ZIKV)-host protein interactions (49, 51) and a newly generated ZIKV-host protein interaction data set were compared by Jianxiong Zeng and colleagues (50). Notably, each study included in the analysis used AP-MS with a different cell type for proteomic analysis and the authors found that most interactions were cell-type specific. However, differences in scoring approaches likely also contribute to the variation, given that virus-host protein interactions

in different cell types but with identical scoring produced more overlap (17, 49). Looking forward, a concerted effort to compare diverse data sets for the same virus will be valuable in improving scoring algorithms and fully understanding caveats to specific experimental approaches.

Virus-Induced Changes in Host Protein Interaction Networks

While we have primarily considered the viral protein the bait, studies focused on host baits in the context of infection can open opportunities for interesting comparative approaches in which the host protein interaction network is altered upon infection or influenced by specific virus mutants (43, 44). These differential protein interaction networks, in which protein interactions are lost or gained during infection (Figure 2a), can reveal mechanisms of replication. One such study focused on cyclic GMP–AMP (cGAMP) synthase (cGAS), a cytoplasmic DNA sensor and a critical member of the innate immune response for DNA and RNA viruses. Krystal Lum and colleagues (43) used quantitative AP-MS to establish the cGAS protein interaction network, including protein interactions that are lost or gained following HSV infection. This led to the identification of OASL as a negative regulator of cGAS activation, thus delineating fundamental cellular regulatory mechanisms important for many viruses. Future work mapping differential protein interaction networks may also help delineate mechanisms of disease, if protein interactions that are lost or gained during infection disrupt the physiological functions of those pathways.

Cross-Virus and -Host Comparisons

Substantial evolutionary insight can be revealed by comparing virus-host protein interactions across multiple virus and/or host species (Figure 2b). For example, virus- and host-specific protein interactions can influence host susceptibility (52), host restriction (53, 54), evasion of host restriction factors (55), and vector competence (56, 57). However, conserved interactions can reveal essential mechanisms of replication for a family of viruses (58) or for a virus with broad host range (40). Using systems biology approaches to perform comprehensive and systematic comparisons of virus-host protein interaction networks across virus and host species has the potential to reveal evolutionary insights on a larger scale. For example, in a recent comparative study, David Gordon and colleagues (59) identified virus-host protein interactions across multiple coronaviruses. These protein interactions were dissected to identify conserved molecular mechanisms of replication, virus-specific dysregulation of the immune response, and potential pan-coronavirus therapeutic targets.

Comparative approaches are especially valuable for viruses that involve host switching for transmission. For example, arthropod-borne viruses typically alternate between a vertebrate and an arthropod host. Because many of the molecular mechanisms of replication are conserved in the vertebrate and arthropod hosts, arthropod-borne viruses must maintain some essential virus-host protein interactions in both hosts. We recently performed comparative mapping of virus-host protein interactions for two flaviviruses, dengue virus and ZIKV, which are transmitted to humans by *Aedes* mosquitoes (49). By identifying flavivirus-human and flavivirus-*Aedes* protein interactions, we identified host- and virus-specific interactions, and interactions conserved across both viruses and both host types. One such conserved interaction, the SEC61 translocon, is critical for flavivirus transmembrane

protein biogenesis and could be targeted pharmacologically to inhibit replication in human and *Aedes* cells (49). Such conserved virus-host protein interactions likely constrain virus evolution due to the need to maintain protein interactions with host protein homologs in two highly divergent host species. Consequently, targeting such conserved protein interactions therapeutically may limit viral escape and represents a potential Achilles' heel for arthropod-borne viruses.

INTEGRATIVE APPROACHES

Integration of systems-level virus-host protein interaction data with complementary approaches is essential to determine how these protein interactions affect virus replication, cell signaling pathways, and viral pathogenesis. Here we review how gene perturbation screens, complementary proteomic approaches, computational predictions of protein structure and interactions, high-throughput model organisms, and patient disease databases can be used to define molecular mechanisms of virus replication and pathogenesis (Figure 3). We also highlight how recent technological advancements and coordinated collaborative efforts have the potential to make this discovery process more efficient.

Gene Perturbation Screens

One of the first questions that arises once a set of virus-host protein interactions is identified is what effect these host factors have on virus replication. Phenotypic gene perturbation screens are the natural next step to answer this question, through either host factor depletion or overexpression (60). While host factor depletion and overexpression are both useful approaches, we focus our discussion on the more common depletion studies. Knockdown [RNA interference or clustered regularly interspaced short palindromic repeats (CRISPR) inference] or knockout (haploid gene trap or CRISPR) strategies can be used for host factor depletion, each with their own advantages and disadvantages. Knockout may result in stronger phenotypes compared to knockdown. However, knockout precludes the study of essential host genes. Over the past 15 years, numerous genome-wide gene perturbation screens have been performed for virus replication phenotypes using knockdown and knockout technologies. These can and have been integrated into virus-host protein interaction networks to relate virus replication phenotypes to molecular mechanisms. For example, in the first comprehensive study of HIV-host protein interactions, Jäger and colleagues (17) found significant enrichment in previously identified HIV host dependency factors derived from four recent genome-wide gene perturbation screens.

As gene perturbation screening capabilities become more widespread, there is value in performing a follow-up gene perturbation screen focused on host factors involved in virus-host protein interactions. These secondary screens have the potential to capture phenotypes that may not have emerged in genome-wide screens, as was the case for virus-host protein interaction studies on influenza A virus (61) and ZIKV (51). Focused follow-up gene perturbation screens also allow for testing additional parameters, such as identifying replication phenotypes across different cell types (59) and viruses (62). More limited screening of hundreds of host factors (instead of thousands) can be used for additional training of virus-host protein interaction scoring algorithms, especially when gold standard

protein interactions (see the section titled Scoring Protein Interaction Data) are limited. This type of integrative innovation was pioneered by Holly Ramage and colleagues (63) for hepatitis C virus (HCV).

Complementary Proteomic Approaches

Additional proteomic data can also be layered on top of virus-host protein interaction data to further dissect molecular mechanisms. These include proteome-wide changes in PTMs and in protein abundance and subcellular localization over the course of infection.**[**AU: Edit OK?**]**

Post-translational modifications.—PTMs of the viral or host proteome can alter protein function important for virus replication and/or the host response (64–67). PTMs such as protein phosphorylation, ubiquitylation, and acetylation can be measured on a proteome-wide scale using specialized enrichment techniques for each PTM (68–71). On their own, PTM proteomic profiling studies have provided insight into mechanisms of virus replication (72–74). However, a few recent studies have combined comprehensive virus-host protein interaction mapping with PTM profiling (51, 75). In a recent study, Alexey Stukalov and colleagues (75) combined systematic AP-MS and PTM profiling of proteome phosphorylation, ubiquitylation, and acetylation to identify molecular mechanisms by which severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) co-opts cellular pathways. The authors used a network diffusion model to integrate the complementary proteomic data to connect physical interactions with downstream pathway regulation based on massive changes in PTMs. This type of analysis measures the distance (number of proteins) between protein interactions and downstream regulatory events, and compares this distance to a randomized data set. A significantly shorter distance in the experimental data set compared to the randomized control indicated that these protein interactions are mechanistically important for downstream effects, which the authors use to guide drug repurposing efforts. Ultimately, these types of integrative approaches will allow researchers to go more rapidly from protein interaction to function, beyond virus replication phenotypes.

Spatio-temporal proteomics.—Protein abundance and subcellular localization are key governing factors of protein function and interaction with other substrates. Virus infection can induce the expression or target degradation of proteins (76, 77). Viruses alter subcellular protein location and reorganize organelle structure to facilitate virus replication (78–80). Therefore, characterizing temporally and spatially resolved changes in protein abundance and connecting them back to virus-host protein interactions can provide key insights into the fundamental mechanisms of replication.

Several elegant studies have related temporal changes in host protein abundance to virus-host protein interactions. One such study by Katie Nightingale and colleagues (81) focused on HCMV, which induces the degradation of many host proteins in an effort to evade host innate immunity. The authors monitored protein abundance over time following HCMV infection and determined proteins that were actively degraded by the lysosome and proteasome using specific inhibitors of these pathways. Using block deletion mutants and AP-MS, the authors were able to identify HCMV protein UL145 to interact with HTLF in

complex with the Cullin 4E3 ligase and target HTLF for degradation, thereby promoting HCMV replication.

Quantitative temporal proteomics can be coupled with biochemical subcellular fractionation approaches to determine how protein subcellular localization is altered by virus infection. Each protein has a unique distribution profile across various fractions that is later used to assign its spatial location. The detailed workflow along with the advantages and the limitations of various subcellular fractionation techniques have been reviewed in detail recently (82–84). In a first-of-its-kind study, Pierre Jean Beltran and colleagues (85) demonstrated how HCMV remodels the architecture of the human proteome over 5 days. In the future, combining this approach with TPCA during infection (47) has the potential to link protein localization and protein interaction data on a large scale.

Protein Structure and Interaction Predictions

Breaking a virus-host protein interaction through protein modification is ultimately one of the most convincing and elegant approaches to demonstrating its importance. Protein truncations and alanine scanning are often used to accomplish this task. Such unbiased approaches have identified viral protein mutants that affect virus replication, antagonism of the host immune response, and pathogenesis through the disruption of the virus-host protein interaction (50, 58, 86–88). Structural information can be helpful to drill down to molecular-level resolution of virus-host protein interactions (89), especially if working with a large virus-host protein interaction data set. However, structures are often not available for the viral protein, host protein, or both. Computational modeling is becoming increasingly established in predicting protein structures and interactions, and holds great promise for the systems-level study of virus-host protein interactions (15, 90).

The major structural prediction tools rely on template-based and ab initio predictions. Template-based approaches rely on proteins with known structures, and related structures are predicted based on suitable homology (91). However, successful ab initio prediction of structures with no minimal similarity to known structures opens many opportunities in protein structure prediction (92). Many prediction tools use a hybrid of template-based and ab initio approaches (93–95). Recent successes incorporating artificial intelligence (AI) have re-energized the field. DeepMind’s AI-driven AlphaFold2 demonstrated outstanding results in the fourteenth Critical Assessment of Structure Prediction (96), a major biennial protein structure prediction conference (97). Structures of protein complexes can also be predicted using AlphaFold-Multimer (98). RoseTTAFold is a similar AI-driven folding algorithm inspired by AlphaFold2 (99). While AlphaFold2 outperforms RoseTTAFold predictions, this increased performance comes with a higher computational cost compared to RoseTTAFold. Whether one uses AlphaFold2 or RoseTTAFold, such structural predictions could be especially valuable for several applications. Structures could be used to predict conservation of the protein interaction across nonmodel host or viral species, and aid in the prediction of viral emergence. These predictions could also be used to refine mutagenesis analysis in a rational manner. It is important to highlight that tools such as AlphaFold-Multimer assume a direct interaction and would interface well with Y2H data sets. However, interactions identified by AP-MS and proximity labeling proteomics may not be direct.

Computational methods can also be used to predict virus-host protein interactions and often rely on structural information to make such predictions. P-HIPSTer (Pathogen-Host Interactome Prediction using STructurE similaRity) is one notable example of this type of computational prediction with experimental validation and cross-comparison with existing experimental data sets (100, 101). There was considerable and often significant overlap between P-HIPSTer predictions and published experimental data sets, although this approach may be better suited for comparison to Y2H data sets because of the direct nature of the interaction. Future iterations of P-HIPSTer that incorporate AlphaFold2 or RoseTTAFold structural predictions could vastly improve virus-host protein interaction predictions.

High-Throughput Modeling of Viral Disease

Molecular-level virus-host protein interactions can cause disease on a macroscopic scale. However, convincingly connecting these two length scales using an infection model of disease requires a permissive animal model capable of generating disease (often a small rodent for viruses that cause human disease) and careful dissection of the virus-host protein interaction with virus and/or host genetics. Jianxiong Zeng and colleagues (50) used systematic AP-MS to identify a critical interaction for ZIKV, which disrupts brain development in utero. Dicer1 was found to interact with ZIKV capsid and was essential for ZIKV replication in neural stem cells. Mutations in ZIKV that disrupted the capsid-Dicer1 interaction resulted in less neuropathogenesis in a Dicer1-dependent manner in a fetal mouse model. While this study is an incredibly elegant example of how ZIKV disrupts brain development, this approach will not be able to efficiently connect virus-host protein interactions to viral pathogenesis at a systems level.

A concerted use of high-throughput animal models will be valuable in making connections to virus-host protein interactions on a larger scale. Model organisms such as *Caenorhabditis elegans* (nematodes), *Drosophila melanogaster* (flies), and *Danio rerio* (zebrafish) have been essential for uncovering fundamental biology related to human disease. These model organisms are known for their low cost, fast reproductive cycles, and facile genetics, making high-throughput screens possible. While infection of high-throughput model organisms has been used to study viral disease (102–104), reductionist approaches involving transgenic expression of single viral proteins can be used to unravel how viral proteins cause disease. Several targeted studies have leveraged transgenic viral protein expression in atypical model organisms to study viral pathogenesis. For example, transgenic zebrafish expressing hepatitis B virus and HCV proteins produced virus-related liver pathologies including the formation of hepatocarcinomas (105–107). Transgenic expression of specific HCMV proteins in flies and zebrafish also disrupted development (108, 109), which is notable given the ability of HCMV to cause congenital birth defects. A study of transgenic flies expressing SARS-CoV-2 protein ORF3a in the central nervous system resulted in phenotypes that could help explain long-term coronavirus disease symptoms such as fatigue, headache, and cognitive impairments (110, 111). Thus, while viral pathogenesis is often complex, some aspects of these complex diseases can be studied using reductionist systems. Systematically connecting these pathologies to specific virus-host protein interactions is critical.

Integrating comprehensive virus-host protein interaction studies with transgenic animal models can bridge this divide more rapidly. For example, in collaboration with Nichole Link and colleagues (49, 112), we used transgenic flies expressing ZIKV NS4A and fly host genetics to demonstrate how NS4A inhibits brain development in an ANKLE2-dependent manner. Given that we also found NS4A to physically interact with ANKLE2 in a large-scale AP-MS study, our work demonstrates how these two approaches can be combined to unravel pathogenic mechanisms with molecular resolution. In another study, Ki-Jun Yoon and colleagues (113) established that transgenic expression of ZIKV NS2A induces defects in cortical neurogenesis. The authors ultimately used AP-MS to identify protein interactions between NS2A and host adherens junctions proteins that are likely responsible for this phenotype. This study is notable because the systematic study of each ZIKV protein in transgenic fetal mice has the advantage of recapitulating human brain development more completely. However, mice do not have the same throughput, making host genetic rescues and systematic transgenic studies for teratogenic large DNA viruses such as HCMV more challenging. Coordinated efforts to systematically identify virus-host protein interactions simultaneously with host phenotypes in high-throughput animal models will be critical to efficiently move from systems to pathogenesis.

Leveraging Similarities with Genetic Disease

The intersection of viral and genetic diseases can also be used to elucidate disease mechanisms underlying both etiologies. For example, the intersection of cancer mutations and virus-host protein interactions for oncogenic viruses can identify new oncogenes. In one of the seminal systems biology studies of virus-host protein interactions, Orit Rosenblatt-Rosen and colleagues (16) explored the hypothesis that genomic variation dysregulates cell division to cause cancer via similar mechanisms as virus-host protein interactions for oncogenic viruses. The authors found significant overlap between host proteins involved in virus-host interactions and those implicated in tumorigenesis in animal models of cancer.

This intersection between viral and genetic disease can also inform on synergies between these similar diseases with distinct etiologies. A recent study by Manon Eckhardt and colleagues (114) focused on virus-host protein interactions for human papillomaviruses (HPVs), associated with cervical and head and neck cancers. The authors integrated the protein interaction data with comprehensive cancer sequencing data from The Cancer Genome Atlas, a comprehensive effort to sequence cancer genomes and characterize tumors at the molecular level (115, 116). This integrative approach identified gene mutations that predisposed individuals to cervical or head and neck cancer following HPV infection. Similarly, our work on ZIKV inhibition of brain development through a virus-host protein interaction (49, 112) was facilitated by the existence of pathogenic gene variants resulting in similar hereditary disease in humans (117, 118) and could relate to host genetic susceptibility to viral microcephaly. For teratogenic viruses such as ZIKV and HCMV, integrating virus-host protein interaction data with rare disease databases, such as the Undiagnosed Diseases Network and Online Mendelian Inheritance in Man (OMIM), will facilitate uncovering disease mechanisms more rapidly.

CONCLUDING REMARKS

Systems biology approaches to identify virus-host protein interactions have uncovered a wealth of information with which to generate hypotheses regarding molecular mechanisms of virus replication and disease. Given the sheer volume of protein interaction data that can now be generated in a short period of time, comparative and integrative systems biology approaches offer an avenue to test these hypotheses more rapidly. In the future, efforts to increase the throughput of hypothesis testing will propel the field forward.

ACKNOWLEDGMENTS

We apologize to colleagues whose work we could not cite due to space constraints. Funding in the Shah lab for research on the systems biology of virus-host protein interactions is provided by the W. M. Keck Foundation.

LITERATURE CITED

1. Fields S, Song O. 1989. A novel genetic system to detect protein–protein interactions. Nature 340(6230):245–46 [PubMed: 2547163]
2. Johnsson N, Varshavsky A. 1994. Split ubiquitin as a sensor of protein interactions *in vivo*. PNAS 91(22):10340–44 [PubMed: 7937952]
3. Iyer K, Bürkle L, Auerbach D, Thaminy S, Dinkel M, et al. 2005. Utilizing the split-ubiquitin membrane yeast two-hybrid system to identify protein-protein interactions of integral membrane proteins. Sci. STKE 2005(275):pl3 [PubMed: 15770033]
4. Yachie N, Petsalaki E, Mellor JC, Weile J, Jacob Y, et al. 2016. Pooled-matrix protein interaction screens using Barcode Fusion Genetics. Mol. Syst. Biol. 12(4):863 [PubMed: 27107012]
5. Trigg SA, Garza RM, MacWilliams A, Nery JR, Bartlett A, et al. 2017. CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. Nat. Methods 14(8):819–25 [PubMed: 28650476]
6. Calderwood MA, Venkatesan K, Xing L, Chase MR, Vazquez A, et al. 2007. Epstein-Barr virus and virus human protein interaction maps. PNAS 104(18):7606–11 [PubMed: 17446270]
7. Lemasson M, Caignard G, Unterfinger Y, Attoui H, Bell-Sakyi L, et al. 2021. Exploration of binary protein–protein interactions between tick-borne flaviviruses and *Ixodes ricinus*. Parasites Vectors 14(1):144 [PubMed: 33676573]
8. Zhang X, Fei D, Sun L, Li M, Ma Y, et al. 2019. Identification of the novel host protein interacting with the structural protein VP1 of Chinese sacbrood virus by yeast two-hybrid screening. Front. Microbiol. 10:2192 [PubMed: 31611854]
9. Shevchenko A, Jensen ON, Podtelejnikov AV, Sagliocco F, Wilm M, et al. 1996. Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. PNAS 93(25):14440–45 [PubMed: 8962070]
10. De Maio FA, Risso G, Iglesias NG, Shah P, Pozzi B, et al. 2016. The dengue virus NS5 protein intrudes in the cellular spliceosome and modulates splicing. PLOS Pathog. 12(8):e1005841 [PubMed: 27575636]
11. Hafirassou ML, Meertens L, Umaña-Diaz C, Labeau A, Dejarnac O, et al. 2017. A global interactome map of the dengue virus NS1 identifies virus restriction and dependency host factors. Cell Rep. 21(13):3900–13 [PubMed: 29281836]
12. Luo Y, Jacobs EY, Greco TM, Mohammed KD, Tong T, et al. 2016. HIV–host interactome revealed directly from infected cells. Nat. Microbiol. 1(7):16068 [PubMed: 27375898]
13. Cristea IM, Rozjabek H, Molloy KR, Karki S, White LL, et al. 2010. Host factors associated with the Sindbis virus RNA-dependent RNA polymerase: role for G3BP1 and G3BP2 in virus replication. J. Virol. 84(13):6720–32 [PubMed: 20392851]
14. Cristea IM, Moorman NJ, Terhune SS, Cuevas CD, O'Keefe ES, et al. 2010. Human cytomegalovirus pUL83 stimulates activity of the viral immediate-early promoter through its interaction with the cellular IFI16 protein. J. Virol. 84(15):7803–14 [PubMed: 20504932]

15. Li T, Chen J, Cristea IM. 2013. Human cytomegalovirus tegument protein pUL83 inhibits IFI16-mediated DNA sensing for immune evasion. Cell Host Microbe 14(5):591–99 [PubMed: 24237704] 15A.Cristea IM, Carroll J-WN, Rout MP, Rice CM, Chait BT, MacDonald MR. 2006. Tracking and Elucidating *Alphavirus* -Host Protein Interactions. Journal of Biological Chemistry. 281(40):30269–78 [PubMed: 16895903] 15B.Lin AE, Greco TM, Döhner K, Sodeik B, Cristea IM. 2013. A Proteomic Perspective of Inbuilt Viral Protein Regulation: pUL46 Tegument Protein is Targeted for Degradation by ICP0 during Herpes Simplex Virus Type 1 Infection*. Molecular & Cellular Proteomics. 12(11):3237–52 [PubMed: 23938468] 15C.Betsinger CN, Jankowski CSR, Hofstadter WA, Federspiel JD, Otter CJ, et al. 2021. The human cytomegalovirus protein pUL13 targets mitochondrial cristae architecture to increase cellular respiration during infection. PNAS. 118(32):15D.Bartuli J, Lorenzi I, Backes S, Grimm C, Fischer U. 2022. A generic protocol for the affinity-purification of native macromolecular complexes from poxvirus-infected cells. STAR Protocols. 3(1):101116 [PubMed: 35118428] 15E.Grimm C, Hillen HS, Bedenk K, Bartuli J, Neyer S, et al. 2019. Structural Basis of Poxvirus Transcription: Vaccinia RNA Polymerase Complexes. Cell. 179(7):1537–1550.e19 [PubMed: 31835032]

16. Rozenblatt-Rosen O, Deo RC, Padi M, Adelmant G, Calderwood MA, et al. 2012. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. Nature 487(7408):491–95 [PubMed: 22810586]
17. Jäger S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, et al. 2012. Global landscape of HIV–human protein complexes. Nature 481(7381):365–70
18. Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, et al. 2012. Viral immune modulators perturb the human molecular network by common and unique strategies. Nature 487(7408):486–90 [PubMed: 22810585]
19. Zhang W, Du J, Evans SL, Yu Y, Yu X-F. 2012. T-cell differentiation factor CBF-β regulates HIV-1 Vif-mediated evasion of host restriction. Nature 481(7381):376–79
20. Jäger S, Kim DY, Hultquist JF, Shindo K, LaRue RS, et al. 2011. Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection. Nature 481(7381):371–75 [PubMed: 22190037]
21. Choi-Rhee E, Schulman H, Cronan JE. 2004. Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. Protein Sci. 13(11):3043–50 [PubMed: 15459338]
22. Roux KJ, Kim DI, Raida M, Burke B. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J. Cell Biol. 196(6):801–10 [PubMed: 22412018]
23. Lam SS, Martell JD, Kamer KJ, Deerinck TJ, Ellisman MH, et al. 2015. Directed evolution of APEX2 for electron microscopy and proximity labeling. Nat. Methods 12(1):51–54 [PubMed: 25419960]
24. Kim DI, Jensen SC, Noble KA, Kc B, Roux KH, et al. 2016. An improved smaller biotin ligase for BioID proximity labeling. Mol. Biol. Cell 27(8):1188–96 [PubMed: 26912792]
25. Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, et al. 2018. Efficient proximity labeling in living cells and organisms with TurboID. Nat. Biotechnol. 36(9):880–87 [PubMed: 30125270]
26. Gaucherand L, Porter BK, Levene RE, Price EL, Schmaling SK, et al. 2019. The influenza A virus endoribonuclease PA-X usurps host mRNA processing machinery to limit host gene expression. Cell Rep. 27(3):776–92.e7 [PubMed: 30995476]
27. Dong S, Wang R, Yu R, Chen B, Si F, et al. 2021. Identification of cellular proteins interacting with PEDV M protein through APEX2 labeling. J. Proteom 240:104191
28. V'kovski P, Gerber M, Kelly J, Pfaender S, Ebert N, et al. 2019. Determination of host proteins composing the microenvironment of coronavirus replicase complexes by proximity-labeling. eLife 8:e42037 [PubMed: 30632963]
29. Nkosi D, Sun L, Duke LC, Patel N, Surapaneni SK, et al. 2020. Epstein-Barr virus LMP1 promotes syntenin-1- and Hrs-induced extracellular vesicle formation for its own secretion to increase cell proliferation and migration. mBio 11(3):e00589–20 [PubMed: 32546618]
30. Boruchowicz H, Hawkins J, Cruz-Palomar K, Lippé R. 2020. The XPO6 exportin mediates herpes simplex virus 1 gM nuclear release late in infection. J. Virol. 94(21):e00753–20 [PubMed: 32817212]

31. Zhang Y, Song G, Lal NK, Nagalakshmi U, Li Y, et al. 2019. TurboID-based proximity labeling reveals that UBR7 is a regulator of N NLR immune receptor-mediated immunity. Nat. Commun. 10(1):3252 [PubMed: 31324801]
32. Liu X, Huuskonen S, Laitinen T, Redchuk T, Bogacheva M, et al. 2021. SARS-CoV-2–host proteome interactions for antiviral drug discovery. Mol. Syst. Biol. 17(11):e10396 [PubMed: 34709727]
33. Kumar A, Salemi M, Bhullar R, Guevara-Plunkett S, Lyu Y, et al. 2021. Proximity biotin labeling reveals Kaposi's sarcoma-associated herpesvirus interferon regulatory factor networks. J. Virol. 95(9):e02049–20 [PubMed: 33597212]
34. Coyaud E, Ranadheera C, Cheng D, Gonçalves J, Dyakov BJA, et al. 2018. Global interactomics uncovers extensive organellar targeting by Zika virus. Mol. Cell. Proteom. 17(11):2242–55
35. Mellacheruvu D, Wright Z, Couzens AL, Lambert J-P, St-Denis NA, et al. 2013. The CRAPome: a contaminant repository for affinity purification–mass spectrometry data. Nat. Methods 10(8):730–36 [PubMed: 23921808]
36. Davis ZH, Verschueren E, Jang GM, Kleffman K, Johnson JR, et al. 2015. Global mapping of herpesvirus-host protein complexes reveals a transcription strategy for late genes. Mol. Cell 57(2):349–60 [PubMed: 25544563]
37. Choi H, Larsen B, Lin Z-Y, Breitkreutz A, Mellacheruvu D, et al. 2011. SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nat. Methods 8(1):70–73 [PubMed: 21131968]
38. Teo G, Koh H, Fermin D, Lambert J-P, Knight JDR, et al. 2016. SAINTq: scoring protein-protein interactions in affinity purification–mass spectrometry experiments with fragment or peptide intensity data. Proteomics 16(15–16):2238–45 [PubMed: 27119218]
39. Kane JR, Stanley DJ, Hultquist JF, Johnson JR, Mietrach N, et al. 2015. Lineage-specific viral hijacking of non-canonical E3 ubiquitin ligase cofactors in the evolution of Vif anti-APOBEC3 activity. Cell Rep. 11(8):1236–50 [PubMed: 25981045]
40. Gestuveo RJ, Royle J, Donald CL, Lamont DJ, Hutchinson EC, et al. 2021. Analysis of Zika virus capsid-*Aedes aegypti* mosquito interactome reveals pro-viral host factors critical for establishing infection. Nat. Commun. 12(1):2766 [PubMed: 33986255]
41. Contu L, Balistreri G, Domanski M, Uldry A-C, Mühlemann O. 2021. Characterisation of the Semliki Forest virus-host cell interactome reveals the viral capsid protein as an inhibitor of nonsense-mediated mRNA decay. PLOS Pathog. 17(5):e1009603 [PubMed: 34019569]
42. DeBlasio SL, Johnson R, Mahoney J, Karasev A, Gray SM, et al. 2015. Insights into the *polerovirus*–plant interactome revealed by coimmunoprecipitation and mass spectrometry. MPMI 28(4):467–81 [PubMed: 25496593]
43. Lum KK, Song B, Federspiel JD, Diner BA, Howard T, Cristea IM. 2018. Interactome and proteome dynamics uncover immune modulatory associations of the pathogen sensing factor cGAS. Cell Syst. 7(6):627–42.e6 [PubMed: 30471916]
44. Hüttenhain R, Xu J, Burton LA, Gordon DE, Hultquist JF, et al. 2019. ARIH2 is a Vif-dependent regulator of CUL5-mediated APOBEC3G degradation in HIV infection. Cell Host Microbe 26(1):86–99.e7 [PubMed: 31253590]
45. Nobre LV, Nightingale K, Ravenhill BJ, Antrobus R, Soday L, et al. 2019. Human cytomegalovirus interactome analysis identifies degradation hubs, domain associations and viral protein functions. eLife 8:e49894 [PubMed: 31873071]
46. Tan CSH, Go KD, Bisteau X, Dai L, Yong CH, et al. 2018. Thermal proximity coaggregation for system-wide profiling of protein complex dynamics in cells. Science 359(6380):1170–77 [PubMed: 29439025]
47. Justice JL, Kennedy MA, Hutton JE, Liu D, Song B, et al. 2021. Systematic profiling of protein complex dynamics reveals DNA-PK phosphorylation of IFI16 en route to herpesvirus immunity. Sci. Adv. 7(25):eabg6680 [PubMed: 34144993]
48. Johnson KL, Qi Z, Yan Z, Wen X, Nguyen TC, et al. 2021. Revealing protein-protein interactions at the transcriptome scale by sequencing. Mol. Cell 81(19):4091–103.e9 [PubMed: 34348091]

49. Shah PS, Link N, Jang GM, Sharp PP, Zhu T, et al. 2018. Comparative flavivirus-host protein interaction mapping reveals mechanisms of dengue and Zika virus pathogenesis. Cell 175(7):1931–45.e18 [PubMed: 30550790]
50. Zeng J, Dong S, Luo Z, Xie X, Fu B, et al. 2020. The Zika virus capsid disrupts corticogenesis by suppressing Dicer activity and miRNA biogenesis. Cell Stem Cell 27(4):618–32.e9 [PubMed: 32763144]
51. Scaturro P, Stukalov A, Haas DA, Cortese M, Draganova K, et al. 2018. An orthogonal proteomic survey uncovers novel Zika virus host factors. Nature 561(7722):253–57 [PubMed: 30177828]
52. Hueffer K, Parker JSL, Weichert WS, Geisel RE, Sgro J-Y, Parrish CR. 2003. The natural host range shift and subsequent evolution of canine parvovirus resulted from virus-specific binding to the canine transferrin receptor. J. Virol. 77(3):1718–26 [PubMed: 12525605]
53. Sawyer SL, Wu LI, Emerman M, Malik HS. 2005. Positive selection of primate TRIM5α identifies a critical species-specific retroviral restriction domain. PNAS 102(8):2832–37 [PubMed: 15689398]
54. Stremlau M, Perron M, Lee M, Li Y, Song B, et al. 2006. Specific recognition and accelerated uncoating of retroviral capsids by the TRIM5α restriction factor. PNAS 103(14):5514–19 [PubMed: 16540544]
55. Park C, Peng C, Rahman MJ, Haller SL, Tazi L, et al. 2021. Orthopoxvirus K3 orthologs show virus- and host-specific inhibition of the antiviral protein kinase PKR. PLOS Pathog. 17(1):e1009183 [PubMed: 33444388]
56. Liang Y, Gao X-W. 2017. The cuticle protein gene MPCP4 of Myzus persicae (Homoptera: Aphididae) plays a critical role in Cucumber mosaic virus acquisition. J. Econ. Entomol. 110(3):848–53 [PubMed: 28334092]
57. Perry KL, Zhang L, Palukaitis P. 1998. Amino acid changes in the coat protein of cucumber mosaic virus differentially affect transmission by the aphids *Myzus persicae* and *Aphis gossypii*. Virology 242(1):204–10 [PubMed: 9501042]
58. Diep J, Ooi YS, Wilkinson AW, Peters CE, Foy E, et al. 2019. Enterovirus pathogenesis requires the host methyltransferase SETD3. Nat. Microbiol. 4(12):2523–37 [PubMed: 31527793]
59. Gordon DE, Hiatt J, Bouhaddou M, Rezelj VV, Ulferts S, et al. 2020. Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. Science 370(6521):eabe9403 [PubMed: 33060197]
60. Ramage H, Cherry S. 2015. Virus-host interactions: from unbiased genetic screens to function. Annu. Rev. Virol. 2:497–524 [PubMed: 26958926]
61. Watanabe T, Kawakami E, Shoemaker JE, Lopes TJS, Matsuoka Y, et al. 2014. Influenza virus-host interactome screen as a platform for antiviral drug development. Cell Host Microbe 16(6):795–805 [PubMed: 25464832]
62. Li M, Johnson JR, Truong B, Kim G, Weinbren N, et al. 2019. Identification of antiviral roles for the exon–junction complex and nonsense-mediated decay in flaviviral infection. Nat. Microbiol. 4(6):985–95 [PubMed: 30833725]
63. Ramage HR, Kumar GR, Verschueren E, Johnson JR, Von Dollen J, et al. 2015. A combined proteomics/genomics approach links hepatitis C virus infection with nonsense-mediated mRNA decay. Mol. Cell 57(2):329–40 [PubMed: 25616068]
64. Lilley CE, Chaurushiya MS, Boutell C, Landry S, Suh J, et al. 2010. A viral E3 ligase targets RNF8 and RNF168 to control histone ubiquitination and DNA damage responses. EMBO J. 29(5):943–55 [PubMed: 20075863]
65. Giraldo MI, Xia H, Aguilera-Aguirre L, Hage A, van Tol S, et al. 2020. Envelope protein ubiquitination drives entry and pathogenesis of Zika virus. Nature 585(7825):414–19 [PubMed: 32641828]
66. Huffmaster NJ, Sollars PJ, Richards AL, Pickard GE, Smith GA. 2015. Dynamic ubiquitination drives herpesvirus neuroinvasion. PNAS 112(41):12818–23 [PubMed: 26407585]
67. Howard TR, Crow MS, Greco TM, Lum KK, Li T, Cristea IM. 2021. The DNA sensor IFIX drives proteome alterations to mobilize nuclear and cytoplasmic antiviral responses, with its acetylation acting as a localization toggle. mSystems 6(3):e0039721 [PubMed: 34156286]

68. Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, et al. 2011. Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol. Cell 44(2):325–40 [PubMed: 21906983]
69. Ficarro SB, McCleland ML, Stukenberg PT, Burke DJ, Ross MM, et al. 2002. Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. Nat. Biotechnol. 20(3):301–5 [PubMed: 11875433]
70. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, et al. 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325(5942):834–40 [PubMed: 19608861]
71. Mertins P, Qiao JW, Patel J, Udeshi ND, Clauser KR, et al. 2013. Integrated proteomic analysis of post-translational modifications by serial enrichment. Nat. Methods 10(7):634–37 [PubMed: 23749302]
72. Bouhaddou M, Memon D, Meyer B, White KM, Rezelj VV, et al. 2020. The global phosphorylation landscape of SARS-CoV-2 infection. Cell 182(3):685–712.e19 [PubMed: 32645325]
73. Ivanov A, Ramanathan P, Parry C, Ilinykh PA, Lin X, et al. 2020. Global phosphoproteomic analysis of Ebola virions reveals a novel role for VP35 phosphorylation-dependent regulation of genome transcription. Cell. Mol. Life Sci. 77(13):2579–603 [PubMed: 31562565]
74. Wojcechowskyj JA, Didigu CA, Lee JY, Parrish NF, Sinha R, et al. 2013. Quantitative phosphoproteomics reveals extensive cellular reprogramming during HIV-1 entry. Cell Host Microbe 13(5):613–23 [PubMed: 23684312]
75. Stukalov A, Girault V, Grass V, Karayel O, Bergant V, et al. 2021. Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. Nature 594(7862):246–52 [PubMed: 33845483]
76. Weekes MP, Tomasec P, Huttlin EL, Fielding CA, Nusinow D, et al. 2014. Quantitative temporal viromics: an approach to investigate host-pathogen interaction. Cell 157(6):1460–72 [PubMed: 24906157]
77. Ersing I, Nobre L, Wang LW, Soday L, Ma Y, et al. 2017. A temporal proteomic map of Epstein-Barr virus lytic replication in B cells. Cell Rep. 19(7):1479–93 [PubMed: 28514666]
78. Ponia SS, Robertson SJ, McNally KL, Subramanian G, Sturdevant GL, et al. 2021. Mitophagy antagonism by ZIKV reveals Ajuba as a regulator of PINK1 signaling, PKR-dependent inflammation, and viral invasion of tissues. Cell Rep. 37(4):109888 [PubMed: 34706234]
79. Heaton NS, Perera R, Berger KL, Khadka S, LaCount DJ, et al. 2010. Dengue virus nonstructural protein 3 redistributes fatty acid synthase to sites of viral replication and increases cellular fatty acid synthesis. PNAS 107(40):17345–50 [PubMed: 20855599]
80. Kim S-J, Syed GH, Khan M, Chiu W-W, Sohail MA, et al. 2014. Hepatitis C virus triggers mitochondrial fission and attenuates apoptosis to promote viral persistence. PNAS 111(17):6413–18 [PubMed: 24733894]
81. Nightingale K, Lin K-M, Ravenhill BJ, Davies C, Nobre L, et al. 2018. High-definition analysis of host protein stability during human cytomegalovirus infection reveals antiviral factors and viral evasion mechanisms. Cell Host Microbe 24(3):447–60.e11 [PubMed: 30122656]
82. Christopher JA, Stadler C, Martin CE, Morgenstern M, Pan Y, et al. 2021. Subcellular proteomics. Nat. Rev. Methods Primers 1:32 [PubMed: 34549195]
83. Lundberg E, Borner GHH. 2019. Spatial proteomics: a powerful discovery tool for cell biology. Nat. Rev. Mol. Cell Biol. 20(5):285–302 [PubMed: 30659282]
84. Borner GHH. 2020. Organellar maps through proteomic profiling—a conceptual guide. Mol. Cell Proteom. 19(7):1076–87
85. Jean Beltran PM, Mathias RA, Cristea IM. 2016. A portrait of the human organelle proteome in space and time during cytomegalovirus infection. Cell Syst. 3(4):361–73.e6 [PubMed: 27641956]
86. Petit MJ, Kenaston MW, Pham OH, Nagainis AA, Fishburn AT, Shah PS. 2021. Nuclear dengue virus NS5 antagonizes expression of PAF1-dependent immune response genes. PLOS Pathog. 17(11):e1010100 [PubMed: 34797876]
87. Greninger AL, Knudsen GM, Betegon M, Burlingame AL, Derisi JL. 2012. The 3A protein from multiple picornaviruses utilizes the Golgi adaptor protein ACBD3 to recruit PI4KIIIβ. J. Virol. 86(7):3605–16 [PubMed: 22258260]

88. Chan YK, Gack MU. 2016. A phosphomimetic-based mechanism of dengue virus to antagonize innate immunity. Nat. Immunol. 17(5):523–30 [PubMed: 26998762]
89. Wang B, Thurmond S, Zhou K, Sánchez-Aparicio MT, Fang J, et al. 2020. Structural basis for STAT2 suppression by flavivirus NS5. Nat. Struct. Mol. Biol. 27(10):875–85 [PubMed: 32778820]
90. Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, et al. 2020. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature 583(7816):459–68 [PubMed: 32353859]
91. Källberg M, Wang H, Wang S, Peng J, Wang Z, et al. 2012. Template-based protein structure modeling using the RaptorX web server. Nat. Protoc. 7(8):1511–22 [PubMed: 22814390]
92. Bonneau R, Tsai J, Ruczinski I, Chivian D, Rohl C, et al. 2001. Rosetta in CASP4: progress in ab initio protein structure prediction. Proteins 5(Suppl.):119–26
93. Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. Nat. Protoc. 5(4):725–38 [PubMed: 20360767]
94. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10(6):845–58 [PubMed: 25950237]
95. Kim DE, Chivian D, Baker D. 2004. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 32(Suppl.):W526–31 [PubMed: 15215442]
96. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596(7873):583–89 [PubMed: 34265844]
97. Kryshtafovych A, Schwede T, Topf M, Fidelis K, Moult J. 2021. Critical assessment of methods of protein structure prediction (CASP)-Round XIV. Proteins 89(12):1607–17 [PubMed: 34533838]
98. Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, et al. 2021. Protein complex prediction with AlphaFold-Multimer. bioRxiv 2021.10.04.463034. 10.1101/2021.10.04.463034
99. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, et al. 2021. Accurate prediction of protein structures and interactions using a three-track neural network. Science 373(6557):871–76 [PubMed: 34282049]
100. Zhang QC, Petrey D, Garzón JI, Deng L, Honig B. 2013. PrePPI: a structure-informed database of protein–protein interactions. Nucleic Acids Res. 41(D1):D828–33 [PubMed: 23193263]
101. Lasso G, Mayer SV, Winkelmann ER, Chu T, Elliot O, et al. 2019. A structure-informed atlas of human-virus interactions. Cell 178(6):1526–41.e16 [PubMed: 31474372]
102. Zhang R, Earnest JT, Kim AS, Winkler ES, Desai P, et al. 2019. Expression of the Mxra8 receptor promotes alphavirus infection and pathogenesis in mice and *Drosophila*. Cell Rep. 28(10):2647–58.e5 [PubMed: 31484075]
103. Liu W-H, Lin Y-L, Wang J-P, Liou W, Hou RF, et al. 2006. Restriction of vaccinia virus replication by a *ced-3* and *ced-4*-dependent pathway in *Caenorhabditis elegans*. PNAS 103(11):4174–79 [PubMed: 16537504]
104. Palha N, Guivel-Benhassine F, Briolat V, Lutfalla G, Sourisseau M, et al. 2013. Real-time whole-body visualization of chikungunya virus infection and host interferon response in zebrafish. PLOS Pathog. 9(9):e1003619 [PubMed: 24039582]
105. Lu J-W, Yang W-Y, Lin Y-M, Jin S-LC, Yuh C-H. 2013. Hepatitis B virus X antigen and aflatoxin B1 synergistically cause hepatitis, steatosis and liver hyperplasia in transgenic zebrafish. Acta Histochem. 115(7):728–39 [PubMed: 23499292]
106. Shieh Y-S, Chang Y-S, Hong J-R, Chen L-J, Jou L-K, et al. 2010. Increase of hepatic fat accumulation by liver specific expression of hepatitis B virus X protein in zebrafish. Biochim. Biophys. Acta 1801(7):721–30 [PubMed: 20416398]
107. Liu W, Chen J-R, Hsu C-H, Li Y-H, Chen Y-M, et al. 2012. A zebrafish model of intrahepatic cholangiocarcinoma by dual expression of hepatitis B virus X and hepatitis C virus core protein in liver. Hepatology 56(6):2268–76 [PubMed: 22729936]
108. Cazorla-Vázquez S, Steingruber M, Marschall M, Engel FB. 2019. Human cytomegaloviral multifunctional protein kinase pUL97 impairs zebrafish embryonic development and increases mortality. Sci. Rep. 9(1):7219 [PubMed: 31076608]
109. Steinberg R, Shemer-Avni Y, Adler N, Neuman-Silberberg S. 2008. Human cytomegalovirus immediate-early-gene expression disrupts embryogenesis in transgenic *Drosophila*. Transgenic Res. 17(1):105–19 [PubMed: 17912601]

110. Crook H, Raza S, Nowell J, Young M, Edison P. 2021. Long covid—mechanisms, risk factors, and management. BMJ 374:n1648 [PubMed: 34312178]

111. Yang S, Tian M, Johnson AN. 2020. SARS-CoV-2 protein ORF3a is pathogenic in Drosophila and causes phenotypes associated with COVID-19 post-viral syndrome. bioRxiv 2020.12.20.423533. 10.1101/2020.12.20.423533

112. Link N, Chung H, Jolly A, Withers M, Tepe B, et al. 2019. Mutations in *ANKLE2*, a ZIKA virus target, disrupt an asymmetric cell division pathway in *Drosophila* neuroblasts to cause microcephaly. Dev. Cell 51(6):713–29.e6 [PubMed: 31735666]

113. Yoon K-J, Song G, Qian X, Pan J, Xu D, et al. 2017. Zika-virus-encoded NS2A disrupts mammalian cortical neurogenesis by degrading adherens junction proteins. Cell Stem Cell 21(3):349–58.e6 [PubMed: 28826723]

114. Eckhardt M, Zhang W, Gross AM, Von Dollen J, Johnson JR, et al. 2018. Multiple routes to oncogenesis are promoted by the human papillomavirus–host protein network. Cancer Discov. 8(11):1474–89 [PubMed: 30209081]

115. Burk RD, Chen Z, Saller C, Tarvin K, Carvalho AL, et al. 2017. Integrated genomic and molecular characterization of cervical cancer. Nature 543(7645):378–84 [PubMed: 28112728]

116. Cancer Genome Atlas Netw. 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature 517(7536):576–82 [PubMed: 25631445]

117. Yamamoto S, Jaiswal M, Charng W-L, Gambin T, Karaca E, et al. 2014. A *Drosophila* genetic resource of mutants to study mechanisms underlying human genetic diseases. Cell 159(1):200–14 [PubMed: 25259927]

118. Shaheen R, Maddirevula S, Ewida N, Alsahli S, Abdel-Salam GMH, et al. 2019. Genomic and phenotypic delineation of congenital microcephaly. Genet. Med. 21(3):545–52 [PubMed: 30214071]

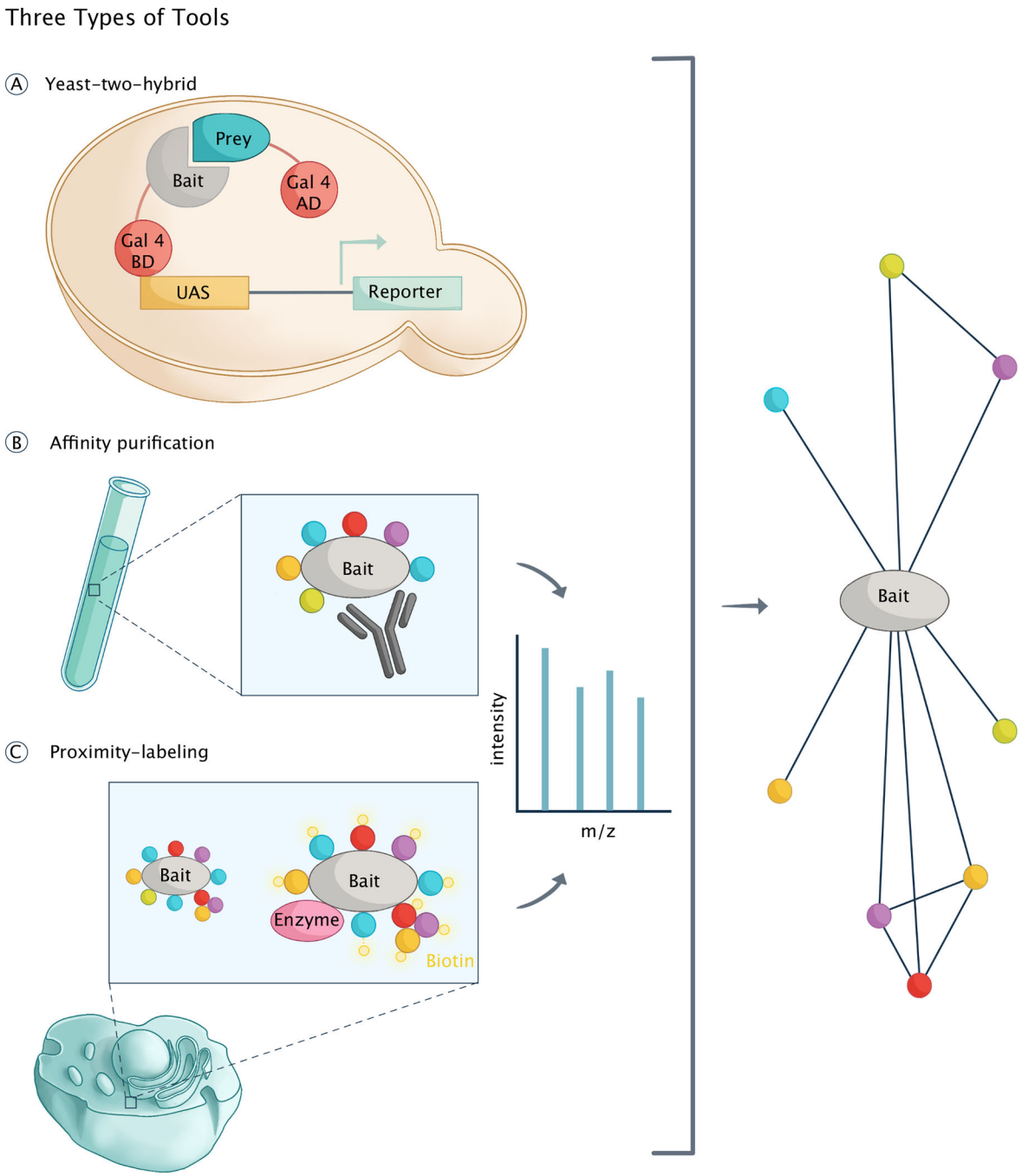

Figure 1.

Common technologies for identifying virus-host protein interactions. (*a*) Yeast two-hybrid (Y2H) is a genetics-based method to identify direct protein interactions through use of selection markers or high-throughput sequencing. The Gal4 binding domain (BD) and activating domain (AD) are fused to the bait and prey, respectively. Successful bait-prey association results in upstream activating sequence (UAS) promoter activation by Gal4 BD and Gal4 AD. (*b*) Affinity purification-mass spectrometry (AP-MS) and (*c*) proximity labeling-mass spectrometry can identify direct and indirect protein interactions. AP-MS

is performed on cellular lysates so multiple subcellular compartments may mix together. Proximity labeling is performed in situ before lysis and preserves some information about subcellular localization. The resulting data from each approach can be visualized as a network of bait and prey, where each prey (*colored circles*) that interacts with the bait (*gray oval*) is connected by a line. The prey part of multi-protein complexes can be shown with lines connecting the prey circles.

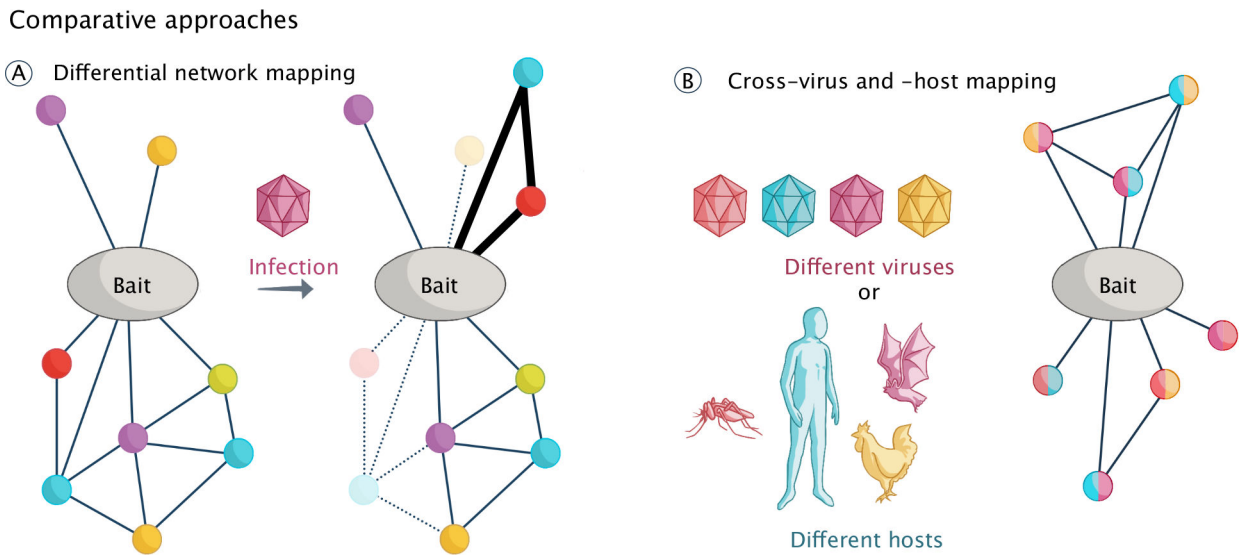

Figure 2.

Comparative mapping of virus-host protein interactions. (*a*) We can identify protein interactions lost or gained during infection and critical to replication or pathogenesis by comparing changes in a host protein interaction network over the course of infection or with different virus mutants. Interactions lost following infection are shown as dashed lines and transparent circles. Interactions gained following infection are shown with thick lines and can be host and/or viral proteins. (*b*) Comparing virus-host protein interactions across different viruses and hosts can reveal highly conserved mechanisms of replication, or virus- or host-specific mechanisms. Interactions shared across two viruses or hosts are shown as split circles with different colors, corresponding to the different virus or host species.

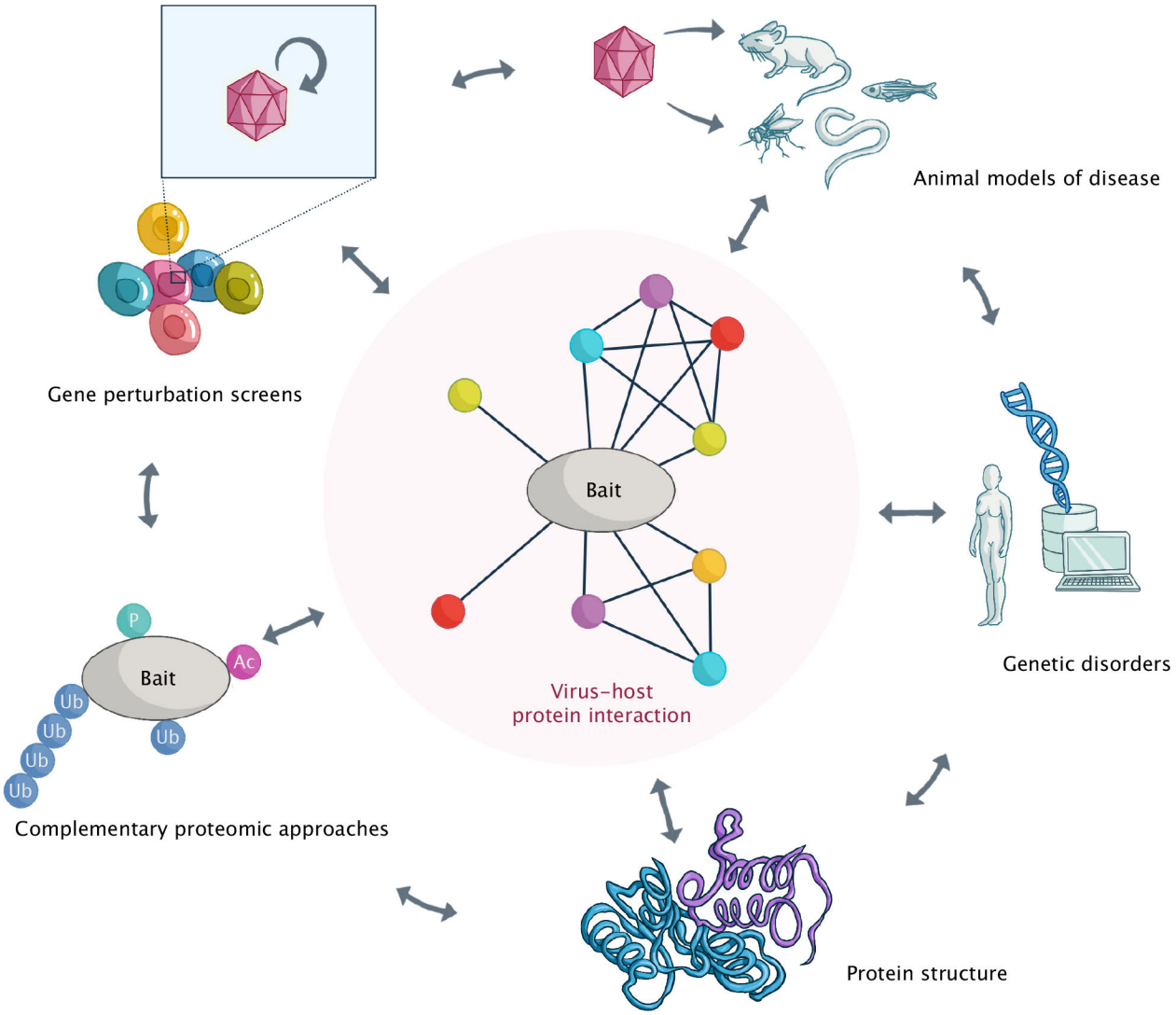

Figure 3.

Integrative approaches. Complementary techniques can be used to go from systems-level virus-host protein interaction data to molecular mechanisms of replication and disease. Effects on virus replication can be assessed in a high-throughput manner using gene perturbation screens. Similarly, effects on cellular pathways and regulatory mechanisms can be assessed on a systems level using complementary proteomic approaches. Structural biology can be used to dissect virus-host protein interactions and disrupt these interactions. Animal models and similarities to human disease can be used to understand viral disease mechanisms. While protein structure and disease mechanisms are typically studied in a targeted manner, advances in computational biology, novel uses of high-throughput animal models, and increasing availability of genome sequencing for human disease can improve the transformation of virus-host protein interaction data into mechanistic insights.