



# Toward a three-dimensional view of protein networks between species

Eric A. Franzosa<sup>1</sup>, Sara Garamszegi<sup>1</sup> and Yu Xia<sup>1,2,3,4\*</sup>

<sup>1</sup> Bioinformatics Program, Boston University, Boston, MA, USA

<sup>2</sup> Department of Chemistry, Boston University, Boston, MA, USA

<sup>3</sup> Department of Biomedical Engineering, Boston University, Boston, MA, USA

<sup>4</sup> Center for Cancer Systems Biology and Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA, USA

## Edited by:

Hironori Sato, National Institute of Infectious Diseases, Japan

## Reviewed by:

Hiroyuki Toh, National Institute of Advanced Industrial Science and Technology, Japan

Jens Von Einem, Institute of Virology, Ulm University Hospital, Germany  
Pascal Braun, Technical University of Munich, Germany

## \*Correspondence:

Yu Xia, Bioinformatics Program, Boston University, Boston, MA 02215, USA.  
e-mail: yuxia@bu.edu

General principles governing biomolecular interactions between species are expected to differ significantly from known principles governing the interactions within species, yet these principles remain poorly understood at the systems level. A key reason for this knowledge gap is the lack of a detailed three-dimensional (3D), atomistic view of biomolecular interaction networks between species. Recent progress in structural biology, systems biology, and computational biology has enabled accurate and large-scale construction of 3D structural models of nodes and edges for protein–protein interaction networks within and between species. The resulting within- and between-species structural interaction networks have provided new biophysical, functional, and evolutionary insights into species interactions and infectious disease. Here, we review the nascent field of between-species structural systems biology, focusing on interactions between host and pathogens such as viruses.

**Keywords: structural systems biology, protein–protein interaction, host–pathogen interaction, bioinformatics and computational biology, network biology**

## INTRODUCTION

Protein–protein interactions (PPIs) can be divided into two fundamentally different classes. The first class of PPIs involves interactions between two proteins encoded within the genome of a single species, where the two proteins cooperate with each other to achieve cellular function in a coordinated fashion. The second class of PPIs involves interactions between two proteins from different species, for example between host proteins and microbial proteins, or between proteins from two different microbial species. These between-species PPIs play key roles in host–microbe and microbe–microbe interactions. Unlike the cooperative PPIs within the host, the interactions between host and microbes are driven by a wide spectrum of co-evolutionary mechanisms, ranging from parasitic to mutualistic (Dethlefsen et al., 2007). General principles of the PPI networks between microbes and their host may differ significantly from known principles governing the cooperative PPI network encoded within the host, yet these principles are not well understood. Here, we review recent progress toward constructing a high-resolution, three-dimensional (3D) structural view of host–pathogen and within-host PPI networks. The resulting host–pathogen and within-host structural interaction networks enable the discovery of new principles of host–pathogen interactions that are otherwise hidden in the binary PPI network. This review focuses on high-throughput mapping and large-scale analysis of host–pathogen PPI networks, which reveal global trends and patterns in host–pathogen interactions that are minimally confounded by investigator biases.

## HOST–PATHOGEN PROTEIN–PROTEIN INTERACTION NETWORKS

The first step toward building host–pathogen structural interaction networks is to map the networks of physical interactions

between host proteins and pathogen proteins. Host–pathogen PPIs have traditionally been studied one at a time. Recently, systems biology approaches have been applied to host–pathogen interaction research. Significant progress has been made in genome-wide mapping of host–pathogen PPI networks (“interactomes”) for many pathogens, especially viruses. Using high-throughput methods such as the yeast two-hybrid system (Fields and Song, 1989) and affinity purification followed by mass spectrometry identification (Rigaut et al., 1999), experimental host–pathogen interactome maps now exist for many viruses (von Schwedler et al., 2003; Uetz et al., 2006; Calderwood et al., 2007; de Chasse et al., 2008; Shapira et al., 2009; Zhang et al., 2009; Khadka et al., 2011; Jager et al., 2012; Pichlmair et al., 2012; Rozenblatt-Rosen et al., 2012). Since viruses are obligate intracellular parasites with small genomes, many, but not all, physical interactions between viral proteins and host proteins have functional importance. Thus, it is essential to complement physical interactome mapping with functional assays that identify host proteins whose perturbation significantly affects viral infection and replication (Brass et al., 2008; König et al., 2008, 2010; Krishnan et al., 2008; Karlas et al., 2010). In addition to host–virus interactome maps, limited host–pathogen interactome data exist for bacterial and eukaryotic pathogens (Dyer et al., 2008, 2010; Mukhtar et al., 2011; Schleker et al., 2012). Since most proteins in bacterial and eukaryotic pathogens do not directly interact with host proteins, a key challenge is to identify pathogen effector proteins that act directly on the host cell to enable infection (Tobe et al., 2006).

Experimental host–pathogen interactome datasets are expected to continue to expand in the near future. The many thousands of experimentally detected host–pathogen PPIs are collected in databases such as VirusMINT (Chatr-aryamontri et al., 2009), VirHostNet (Navratil et al., 2009), IntAct (Aranda et al., 2010),

PIG (Driscoll et al., 2009), and NCBI HIV-1 protein interaction database (Fu et al., 2009). These databases typically rely heavily on manual curation to maintain standards of quality, and there is a great need to complement manual curation with automated literature mining of host–pathogen PPIs (Thieu et al., 2012).

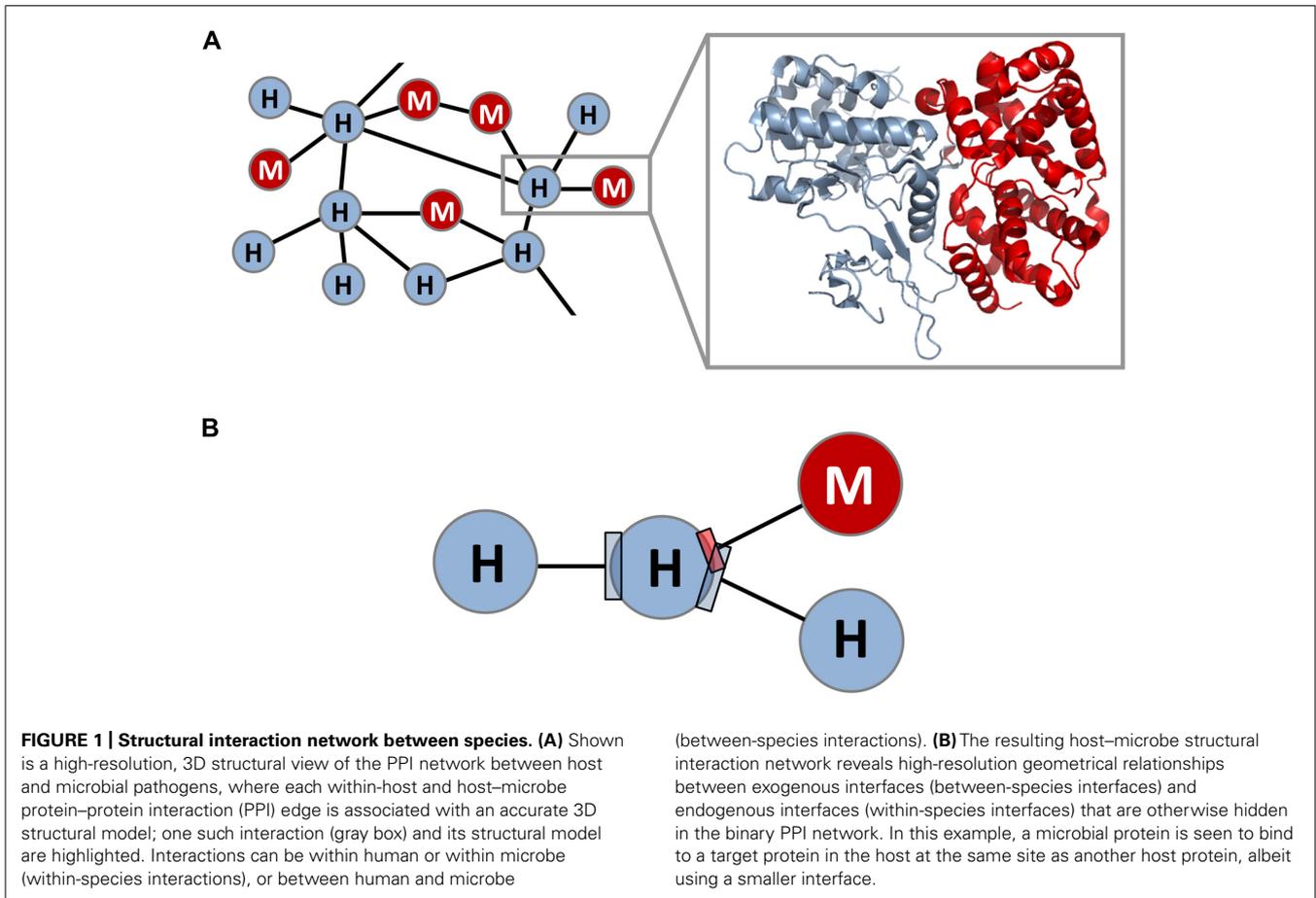
Because of the challenges associated with experimental determination of host–pathogen PPIs, it is desirable to develop computational methods to predict host–pathogen PPIs. Prediction of host–pathogen PPIs is usually based on sequence homology with known PPIs (Uetz et al., 2006; Davis et al., 2007; Doolittle and Gomez, 2011; Wuchty, 2011), the presence of known or predicted interacting domain pairs (Dyer et al., 2007), as well as the presence of other predictive sequence and functional features (Tastan et al., 2009; Qi et al., 2010; Dyer et al., 2011). Computational predictions of host–pathogen PPIs are most effective as a means to prioritize subsequent experimental validations, which are often time-consuming (Uetz et al., 2006). Other areas where computational methods play an increasingly important role include genomic data integration of diverse host–pathogen physical, genetic, and functional interactions (Shapira et al., 2009; Konig et al., 2010; Rozenblatt-Rosen et al., 2012), and network-based prediction of host proteins important in host–pathogen interaction (Navratil et al., 2010; Murali et al., 2011).

Experimental host–pathogen PPI networks are useful in many ways. They not only help generate hypotheses regarding the function of specific pathogen proteins and the biology of specific pathogens, but also provide insights into principles governing host–pathogen interactions at the systems level. Global analyses of host–pathogen PPI networks have revealed that viruses and other microbial pathogens tend to interact with host proteins that are hubs (i.e., proteins with many interaction partners in the host network) and bottlenecks (i.e., proteins whose removal would disrupt many shortest paths in the host network; Calderwood et al., 2007; de Chassey et al., 2008; Dyer et al., 2008; Wuchty et al., 2010; Pichlmair et al., 2012). Host proteins that interact with pathogens tend to be conserved among closely related species (Jager et al., 2012; Pichlmair et al., 2012), although many of them are also under positive selection (Bozek and Lengauer, 2010). Host proteins that interact with pathogens tend to form densely connected network modules by clustering into biological pathways and physical complexes (Dyer et al., 2008; Bushman et al., 2009; MacPherson et al., 2010). In addition, host–pathogen PPI networks are enriched for certain network motifs (e.g., mutual inhibition; van Dijk et al., 2010). Furthermore, pathogens tend to target host proteins involved in common biological processes essential to pathogen infection and replication in general, such as host defense and immune response (Dyer et al., 2008; Pichlmair et al., 2012), often through convergent evolution (Mukhtar et al., 2011). At the same time, different classes of pathogens (e.g., DNA viruses versus RNA viruses, or viruses versus bacteria) also target distinct host pathways due to class-specific differences in infection and replication mechanisms (Durmus Tekir et al., 2012; Pichlmair et al., 2012). Finally, host proteins targeted by pathogens tend to be in network proximity to other proteins implicated in diseases associated with pathogen infections (Navratil et al., 2011; Gulbahce et al., 2012). It is clear that much can be learned by taking a global and network perspective on host–pathogen interactions.

## HOST–PATHOGEN STRUCTURAL INTERACTION NETWORKS

The mapping of host–pathogen PPI networks lays the foundation for and constitutes the first step toward constructing host–pathogen structural interaction networks. Despite experimental and computational advances in the global analysis of host–pathogen PPI networks, the utility of PPI networks is ultimately limited by their low-resolution nature (i.e., proteins represented as nodes and PPIs represented as edges). A high-resolution view of the host–pathogen PPI network can be achieved by building accurate 3D structural models for nodes and edges in the network (**Figure 1A**). Is it feasible to construct such a host–pathogen structural interaction network in a global and accurate way? And if so, does this 3D structural view provide new insights into host–pathogen interactions that are not apparent in the binary PPI network?

Although the 3D structure of proteins and PPIs can in principle be predicted from sequence without resorting to homology [using template-free structure prediction (Moult, 2005) and macromolecular docking (Gray, 2006)], in practice homology modeling remains the most successful and reliable 3D structure prediction method on a genomic scale for both proteins and PPIs (Marti-Renom et al., 2000; Russell et al., 2004). To build a homology model for a query protein or a query pair of interacting proteins, the query protein or protein pair is searched against a template library consisting of proteins or PPIs of known 3D structure deposited in the Protein Data Bank (PDB; Berman et al., 2000). The most significantly matched 3D template is then used to construct a homology model for the query protein or PPI. Despite the obvious limitations that good homology models cannot be built for proteins with entirely new folds or PPIs with entirely new modes of interaction, and that the conformation of proteins and PPIs is not always conserved during evolution, homology modeling has been highly successful in practice, thanks to major advances in structural biology and computational biology. Proteins and PPIs are composed of a limited number of domains and domain–domain interactions (Chothia, 1992; Aloy and Russell, 2004), and certain domains and domain–domain interactions are significantly overrepresented in proteomes and interactomes (Qian et al., 2001). Thus, homology models for many proteins and PPIs can be built based on a relatively small number of representative domains and domain–domain interactions of known 3D structure, stored in databases such as SUPERFAMILY (Madera et al., 2004), iPFam (Finn et al., 2005), and 3did (Stein et al., 2005). Indeed, it is estimated that ~60% of all query proteins share significant sequence similarity with at least one template protein of known 3D structure (Madera et al., 2004). For the vast majority of these cases, the query protein shares significant structural similarity with the template protein, an accurate sequence alignment can be constructed, and an accurate homology model (~3 Å RMSD) can be built for at least a part of the query protein (typically a domain; Marti-Renom et al., 2000; Dalton and Jackson, 2007). Compared to homology modeling of single proteins, the coverage of accurate homology models for within-species PPIs is smaller but still considerable (~20%; Kim et al., 2006). Indeed, it was recently argued that 3D templates exist for most known within-species PPIs, provided that good homology models can be built for the protein components (Kundrotas et al., 2012). The coverage of accurate template-based



models for PPIs can be further improved by identifying additional 3D templates that are structurally similar to the query proteins in the absence of sequence similarity (Zhang et al., 2012).

Homology modeling has been successfully used to construct within-species structural interaction networks, where 3D structural models are built for known within-species PPIs (Aloy et al., 2004; Kim et al., 2006). Despite the caveat that 3D homology models are biased toward soluble, stable, and structurally well-ordered proteins and PPIs, structural interaction networks can be viewed as high-quality subsets of binary PPI networks with much higher spatial resolution. Computational analyses of the within-species structural interaction networks have provided significant insights into a wide range of topics including biophysics, evolution, disease biology, and drug design (Kim et al., 2006, 2008; Franzosa and Xia, 2008, 2009; Kar et al., 2009; Xie et al., 2011; Wang et al., 2012). Such structural systems biology approaches are highly valuable as a unifying framework that integrates molecular biophysics with cell systems biology.

Most recently, structural systems biology was applied to between-species interactions, and an integrated map of human–virus and within-human structural interaction networks was constructed (Franzosa and Xia, 2011). The structural interaction networks consist of 53 human–virus PPIs and >3,000 human–human PPIs in the form of either experimental 3D structures or homology models. Here, instead of predicting new host–pathogen

PPIs (Davis et al., 2007), homology modeling is used to annotate known host–pathogen PPIs with 3D structural information, thus providing a structural map of the binary PPI network in much higher spatial resolution. For example, the binary PPI network indicates that the human CDK6 protein interacts with both human proteins and the cyclin D homolog protein from herpesvirus. The structural interaction network further reveals that these interactions largely occur at two distinct, non-overlapping interfaces on the human CDK6 protein: one interface mediating the interactions with the viral protein as well as the human cyclin D protein, and a second interface mediating the interactions with various human CDK inhibitor proteins (Russo et al., 1998; Pratt et al., 2006). Such a high-resolution map enables the detailed analysis of the geometrical properties and relationships of human–virus PPI interfaces (exogenous interfaces) and human–human PPI interfaces (endogenous interfaces) that is otherwise inaccessible in the binary PPI network (Figure 1B; Franzosa and Xia, 2011). For example, although binary PPI network analysis revealed that viral proteins tend to interact with host protein hubs participating in many endogenous interactions, the precise spatial relationships among these exogenous and endogenous interactions are not known. On the other hand, structural interaction analysis further revealed that exogenous interfaces, although smaller in size, tend to overlap significantly with and mimic endogenous interfaces, often in the absence of sequence or structural

similarity. In addition, the endogenous interfaces that are mimicked by viral proteins tend to participate in multiple endogenous interactions which are transient and regulatory in nature. A case in point is the interaction between the UL36 protein from the HSV-1 virus and the human ubiquitin protein, an important regulator of protein function and cell behavior (Schlieker et al., 2007). The endogenous interface of the human ubiquitin protein mimicked by the virus mediates as many as 30 interactions with other human proteins. On average an endogenous interface mimicked by virus mediates more than three interactions with other human proteins in the structural interaction network, whereas a generic endogenous interface only mediates  $\sim 1.5$  interactions with other human proteins. These observations demonstrate that viral proteins tend to mimic and hijack high-level regulatory components of the host cellular circuitry, by efficiently binding to existing endogenous interfaces rather than creating entirely new interfaces. Furthermore, endogenous interfaces mimicked by viral proteins tend to evolve more quickly than other endogenous interfaces, suggesting an evolutionary “arms race” between host and pathogen. Overall, 3D structural analysis revealed, in a systematic and statistically rigorous way, distinct principles governing antagonism versus cooperation in host–pathogen and within-host PPI networks (Franzosa and Xia, 2011).

Protein–protein interactions can be divided into two classes: the first class involves PPIs mediated by interactions between two globular domains, and the second class involves PPIs mediated by short linear motifs interacting with globular domains. Both classes are important mediators of host–pathogen interactions (Davey et al., 2011; Franzosa and Xia, 2011). A recent survey revealed extensive mimicry of host short linear motifs by viruses (Davey et al., 2011). Viral mimicry of host linear motifs was found for 52 of the  $\sim 150$  motif classes in the Eukaryotic Linear Motif (ELM) database (Gould et al., 2010), 13 of which have solved 3D structures involving viral motifs in complex with their host targets. For example, there are many cases of viral proteins targeting the SH3, SH2, or PDZ domains of host proteins using mimicked motifs. These observations are in agreement with the requirements for viral proteins to extensively hijack and manipulate diverse host proteins and pathways, despite the severe spatial constraints imposed by their small genomes (Davey et al., 2011). These motifs tend to cluster into hotspots in the viral genome (Sarmady et al., 2011), and they may be important determinants of virulence (Yang, 2012). While motifs play an important role in the biology of viruses and viruses use motifs extensively, it is not known if viruses use motifs more

often than the host (Davey et al., 2011). These findings collectively highlight the feasibility and importance of structural systems biology in host–pathogen interaction research.

## CONCLUSION

Despite being a relatively new field, between-species structural systems biology has already provided major insights into species interactions and infectious disease. We expect to see rapid growth in between-species structural systems biology over the next few years on the following fronts. First, host–pathogen physical, genetic, and functional interaction datasets will continue to accumulate for more pathogens, and with higher coverage and accuracy. The impact of these interactions on host and pathogen physiology will continue to be systematically evaluated. In addition to interaction data, small-scale experiments and large-scale technologies such as genome-wide association studies (Khor and Hibberd, 2012) have generated large amounts of data describing mutations that affect host–pathogen interaction and pathogenicity. A key computational challenge is the development of unified, predictive models of how host and pathogens interact through integration of these datasets. Second, the success of homology modeling depends critically on the availability of 3D structural templates for representative proteins and PPIs solved by experimental structural biologists. The power of homology modeling is especially limited for fast-evolving pathogens such as viruses, where experimental structural biology plays a central role. It is encouraging that the number of 3D structures of human–virus PPIs have doubled in the past 5 years (Franzosa and Xia, 2012), and we expect a significant expansion in the number of 3D structures for host–pathogen PPIs in the next few years. Structural genomics has been highly successful by focusing primarily on structure determination of single proteins (Chandonia and Brenner, 2006). It will be fascinating to investigate if high-throughput structural biology can be applied to within- and between-species PPIs as well. Finally, new methods will be developed to integrate interaction datasets with 3D structure datasets. Computational analysis of the resulting structural interaction networks will uncover new system-level insights into host–pathogen interactions.

## ACKNOWLEDGMENTS

This work was supported by a grant from the National Science Foundation (CCF-1219007) to Yu Xia. Sara Garamszegi was supported by a National Science Foundation Graduate Research Fellowship (DGE-0741448).

## REFERENCES

- Aloy, P., Bottcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., et al. (2004). Structure-based assembly of protein complexes in yeast. *Science* 303, 2026–2029.
- Aloy, P., and Russell, R. B. (2004). Ten thousand interactions for the molecular biologist. *Nat. Biotechnol.* 22, 1317–1321.
- Aranda, B., Achuthan, P., Alam-Faruque, Y., Armean, I., Bridge, A., Derow, C., et al. (2010). The IntAct molecular interaction database in 2010. *Nucleic Acids Res.* 38, D525–D531.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242.
- Bozek, K., and Lengauer, T. (2010). Positive selection of HIV host factors and the evolution of lentivirus genes. *BMC Evol. Biol.* 10:186. doi: 10.1186/1471-2148-10-186
- Brass, A. L., Dykxhoorn, D. M., Benita, Y., Yan, N., Engelman, A., Xavier, R. J., et al. (2008). Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 319, 921–926.
- Bushman, F. D., Malani, N., Fernandes, J., D’Orso, I., Cagney, G., Diamond, T. L., et al. (2009). Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathog.* 5:e1000437. doi: 10.1371/journal.ppat.1000437
- Calderwood, M. A., Venkatesan, K., Xing, L., Chase, M. R., Vazquez, A., Holthaus, A. M., et al. (2007). Epstein–Barr virus and virus human protein interaction maps. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7606–7611.
- Chandonia, J. M., and Brenner, S. E. (2006). The impact of structural genomics: expectations and outcomes. *Science* 311, 347–351.
- Chatr-aryamontri, A., Ceol, A., Peluso, D., Nardozza, A., Panni, S., Sacco, F., et al. (2009). VirusMINT: a viral protein interaction database. *Nucleic Acids Res.* 37, D669–D673.
- Chothia, C. (1992). Proteins. One thousand families for the molecular biologist. *Nature* 357, 543–544.

- Dalton, J. A., and Jackson, R. M. (2007). An evaluation of automated homology modelling methods at low target template sequence similarity. *Bioinformatics* 23, 1901–1908.
- Davey, N. E., Trave, G., and Gibson, T. J. (2011). How viruses hijack cell regulation. *Trends Biochem. Sci.* 36, 159–169.
- Davis, F. P., Barkan, D. T., Eswar, N., Mckerrow, J. H., and Sali, A. (2007). Host pathogen protein interactions predicted by comparative modeling. *Protein Sci.* 16, 2585–2596.
- de Chassey, B., Navratil, V., Tafforeau, L., Hiet, M. S., Aublin-Gex, A., Agaugue, S., et al. (2008). Hepatitis C virus infection protein network. *Mol. Syst. Biol.* 4, 230.
- Dethlefsen, L., Mcfall-Ngai, M., and Relman, D. A. (2007). An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature* 449, 811–818.
- Doolittle, J. M., and Gomez, S. M. (2011). Mapping protein interactions between Dengue virus and its human and insect hosts. *PLoS Negl. Trop. Dis.* 5:e954. doi: 10.1371/journal.pntd.0000954
- Driscoll, T., Dyer, M. D., Murali, T. M., and Sobral, B. W. (2009). PIG – the pathogen interaction gateway. *Nucleic Acids Res.* 37, D647–D650.
- Durmus Tekir, S., Cakir, T., and Ulgen, K. O. (2012). Infection strategies of bacterial and viral pathogens through pathogen–human protein–protein interactions. *Front. Microbiol.* 3:46. doi: 10.3389/fmicb.2012.00046
- Dyer, M. D., Murali, T. M., and Sobral, B. W. (2007). Computational prediction of host–pathogen protein–protein interactions. *Bioinformatics* 23, i159–i166.
- Dyer, M. D., Murali, T. M., and Sobral, B. W. (2008). The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog.* 4:e32. doi: 10.1371/journal.ppat.0040032
- Dyer, M. D., Murali, T. M., and Sobral, B. W. (2011). Supervised learning and prediction of physical interactions between human and HIV proteins. *Infect. Genet. Evol.* 11, 917–923.
- Dyer, M. D., Neff, C., Dufford, M., Rivera, C. G., Shattuck, D., Bassaganya-Riera, J., et al. (2010). The human-bacterial pathogen protein interaction networks of *Bacillus anthracis*, *Francisella tularensis*, and *Yersinia pestis*. *PLoS ONE* 5:e12089. doi: 10.1371/journal.pone.0012089
- Fields, S., and Song, O. (1989). A novel genetic system to detect protein–protein interactions. *Nature* 340, 245–246.
- Finn, R. D., Marshall, M., and Bateman, A. (2005). iPfam: visualization of protein–protein interactions in PDB at domain and amino acid resolutions. *Bioinformatics* 21, 410–412.
- Franzosa, E. A., and Xia, Y. (2008). Structural perspectives on protein evolution. *Annu. Rep. Comput. Chem.* 4, 3–21.
- Franzosa, E. A., and Xia, Y. (2009). Structural determinants of protein evolution are context-sensitive at the residue level. *Mol. Biol. Evol.* 26, 2387–2395.
- Franzosa, E. A., and Xia, Y. (2011). Structural principles within the human–virus protein–protein interaction network. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10538–10543.
- Franzosa, E. A., and Xia, Y. (2012). Structural models for host–pathogen protein–protein interactions: assessing coverage and bias. *Pac. Symp. Biocomput.* 287–298.
- Fu, W., Sanders-Bear, B. E., Katz, K. S., Maglott, D. R., Pruitt, K. D., and Ptak, R. G. (2009). Human immunodeficiency virus type 1, human protein interaction database at NCBI. *Nucleic Acids Res.* 37, D417–D422.
- Gould, C. M., Diella, F., Via, A., Puntervoll, P., Gemund, C., Chabanis-Davidson, S., et al. (2010). ELM: the status of the 2010 eukaryotic linear motif resource. *Nucleic Acids Res.* 38, D167–D180.
- Gray, J. J. (2006). High-resolution protein–protein docking. *Curr. Opin. Struct. Biol.* 16, 183–193.
- Gulbahce, N., Yan, H., Dricot, A., Padi, M., Byrdsong, D., Franchi, R., et al. (2012). Viral perturbations of host networks reflect disease etiology. *PLoS Comput. Biol.* 8:e1002531. doi: 10.1371/journal.pcbi.1002531
- Jager, S., Cimermancic, P., Gulbahce, N., Johnson, J. R., McGovern, K. E., Clarke, S. C., et al. (2012). Global landscape of HIV-human protein complexes. *Nature* 481, 365–370.
- Kar, G., Gursoy, A., and Keskin, O. (2009). Human cancer protein–protein interaction network: a structural perspective. *PLoS Comput. Biol.* 5:e1000601. doi: 10.1371/journal.pcbi.1000601
- Karlas, A., Machuy, N., Shin, Y., Pleissner, K. P., Artarini, A., Heuer, D., et al. (2010). Genome-wide RNAi screen identifies human host factors crucial for influenza virus replication. *Nature* 463, 818–822.
- Khadka, S., Vangeloff, A. D., Zhang, C., Siddavatam, P., Heaton, N. S., Wang, L., et al. (2011). A physical interaction network of dengue virus and human proteins. *Mol. Cell. Proteomics* 10, M111.012187.
- Khor, C. C., and Hibberd, M. L. (2012). Host–pathogen interactions revealed by human genome-wide surveys. *Trends Genet.* 28, 233–243.
- Kim, P. M., Lu, L. J., Xia, Y., and Gerstein, M. B. (2006). Relating three-dimensional structures to protein networks provides evolutionary insights. *Science* 314, 1938–1941.
- Kim, P. M., Sboner, A., Xia, Y., and Gerstein, M. (2008). The role of disorder in interaction networks: a structural analysis. *Mol. Syst. Biol.* 4, 179.
- Konig, R., Stertz, S., Zhou, Y., Inoue, A., Hoffmann, H. H., Bhattacharyya, S., et al. (2010). Human host factors required for influenza virus replication. *Nature* 463, 813–817.
- Konig, R., Zhou, Y., Elleder, D., Diamond, T. L., Bonamy, G. M., Ireland, J. T., et al. (2008). Global analysis of host–pathogen interactions that regulate early-stage HIV-1 replication. *Cell* 135, 49–60.
- Krishnan, M. N., Ng, A., Sukumaran, B., Gilfoy, F. D., Uchil, P. D., Sultana, H., et al. (2008). RNA interference screen for human genes associated with West Nile virus infection. *Nature* 455, 242–245.
- Kundrotas, P. J., Zhu, Z., Janin, J., and Vakser, I. A. (2012). Templates are available to model nearly all complexes of structurally characterized proteins. *Proc. Natl. Acad. Sci. U.S.A.* 109, 9438–9441.
- MacPherson, I., Dickerson, J. E., Pinney, J. W., and Robertson, D. L. (2010). Patterns of HIV-1 protein interaction identify perturbed host-cellular subsystems. *PLoS Comput. Biol.* 6:e1000863. doi: 10.1371/journal.pcbi.1000863
- Madera, M., Vogel, C., Kummerfeld, S. K., Chothia, C., and Gough, J. (2004). The SUPERFAMILY database in 2004: additions and improvements. *Nucleic Acids Res.* 32, D235–D239.
- Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29, 291–325.
- Moult, J. (2005). A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction. *Curr. Opin. Struct. Biol.* 15, 285–289.
- Mukhtar, M. S., Carvunis, A. R., Dreze, M., Epple, P., Steinbrenner, J., Moore, J., et al. (2011). Independently evolved virulence effectors converge onto hubs in a plant immune system network. *Science* 333, 596–601.
- Murali, T. M., Dyer, M. D., Badger, D., Tyler, B. M., and Katze, M. G. (2011). Network-based prediction and analysis of HIV dependency factors. *PLoS Comput. Biol.* 7:e1002164. doi: 10.1371/journal.pcbi.1002164
- Navratil, V., De Chassey, B., Combe, C. R., and Lotteau, V. (2011). When the human viral infectome and disease networks collide: towards a systems biology platform for the aetiology of human diseases. *BMC Syst. Biol.* 5:13. doi: 10.1186/1752-0509-5-13
- Navratil, V., De Chassey, B., Meyniel, L., Delmotte, S., Gautier, C., Andre, P., et al. (2009). VirHostNet: a knowledge base for the management and the analysis of proteome-wide virus–host interaction networks. *Nucleic Acids Res.* 37, D661–D668.
- Navratil, V., De Chassey, B., Meyniel, L., Pradezynski, F., Andre, P., Rabourdin-Combe, C., et al. (2010). System-level comparison of protein–protein interactions between viruses and the human type I interferon system network. *J. Proteome Res.* 9, 3527–3536.
- Pichlmair, A., Kandasamy, K., Alvisi, G., Mulhern, O., Sacco, R., Habjan, M., et al. (2012). Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature* 487, 486–490.
- Pratt, D. J., Bentley, J., Jewsbury, P., Boyle, F. T., Endicott, J. A., and Noble, M. E. (2006). Dissecting the determinants of cyclin-dependent kinase 2 and cyclin-dependent kinase 4 inhibitor selectivity. *J. Med. Chem.* 49, 5470–5477.
- Qi, Y., Tastan, O., Carbonell, J. G., Klein-Seetharaman, J., and Weston, J. (2010). Semi-supervised multi-task learning for predicting interactions between HIV-1 and human proteins. *Bioinformatics* 26, i645–i652.
- Qian, J., Luscombe, N. M., and Gerstein, M. (2001). Protein family and fold occurrence in genomes: power-law behaviour and evolutionary model. *J. Mol. Biol.* 313, 673–681.
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030–1032.
- Rozenblatt-Rosen, O., Deo, R. C., Padi, M., Adelman, G., Calderwood, M. A., Rolland, T., et al. (2012). Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* 487, 491–495.
- Russell, R. B., Alber, F., Aloy, P., Davis, F. P., Korkin, D., Pichaud,

- M., et al. (2004). A structural perspective on protein–protein interactions. *Curr. Opin. Struct. Biol.* 14, 313–324.
- Russo, A. A., Tong, L., Lee, J. O., Jeffrey, P. D., and Pavletich, N. P. (1998). Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16INK4a. *Nature* 395, 237–243.
- Sarmady, M., Dampier, W., and Tozeren, A. (2011). HIV protein sequence hotspots for crosstalk with host hub proteins. *PLoS ONE* 6:e23293. doi: 10.1371/journal.pone.0023293
- Schleker, S., Sun, J., Raghavan, B., Srnec, M., Muller, N., Koepfinger, M., et al. (2012). The current *Salmonella*–host interactome. *Proteomics Clin. Appl.* 6, 117–133.
- Schlieker, C., Weihofen, W. A., Frijns, E., Kattenhorn, L. M., Gaudet, R., and Ploegh, H. L. (2007). Structure of a herpesvirus-encoded cysteine protease reveals a unique class of deubiquitinating enzymes. *Mol. Cell* 25, 677–687.
- Shapira, S. D., Gat-Viks, I., Shum, B. O., Dricot, A., De Grace, M. M., Wu, L., et al. (2009). A physical and regulatory map of host–influenza interactions reveals pathways in H1N1 infection. *Cell* 139, 1255–1267.
- Stein, A., Russell, R. B., and Aloy, P. (2005). 3did: interacting protein domains of known three-dimensional structure. *Nucleic Acids Res.* 33, D413–D417.
- Tastan, O., Qi, Y., Carbonell, J. G., and Klein-Seetharaman, J. (2009). Prediction of interactions between HIV-1 and human proteins by information integration. *Pac. Symp. Biocomput.* 516–527.
- Thieu, T., Joshi, S., Warren, S., and Korkin, D. (2012). Literature mining of host–pathogen interactions: comparing feature-based supervised learning and language-based approaches. *Bioinformatics* 28, 867–875.
- Tobe, T., Beatson, S. A., Taniguchi, H., Abe, H., Bailey, C. M., Fivian, A., et al. (2006). An extensive repertoire of type III secretion effectors in *Escherichia coli* O157 and the role of lambdoid phages in their dissemination. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14941–14946.
- Uetz, P., Dong, Y. A., Zeretzke, C., Atzler, C., Baiker, A., Berger, B., et al. (2006). Herpesviral protein networks and their interaction with the human proteome. *Science* 311, 239–242.
- van Dijk, D., Ertaylan, G., Boucher, C. A. B., and Sloot, P. M. A. (2010). Identifying potential survival strategies of HIV-1 through virus–host protein interaction networks. *BMC Syst. Biol.* 4:96. doi: 10.1186/1752-0509-4-96
- von Schwedler, U. K., Stuchell, M., Muller, B., Ward, D. M., Chung, H. Y., Morita, E., et al. (2003). The protein network of HIV budding. *Cell* 114, 701–713.
- Wang, X., Wei, X., Thijssen, B., Das, J., Lipkin, S. M., and Yu, H. (2012). Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat. Biotechnol.* 30, 159–164.
- Wuchty, S. (2011). Computational prediction of host–parasite protein interactions between *P. falciparum* and *H. sapiens*. *PLoS ONE* 6:e26960. doi: 10.1371/journal.pone.0026960
- Wuchty, S., Siwo, G., and Ferdig, M. T. (2010). Viral organization of human proteins. *PLoS ONE* 5:e11796. doi: 10.1371/journal.pone.0011796
- Xie, L., Xie, L., and Bourne, P. E. (2011). Structure-based systems biology for analyzing off-target binding. *Curr. Opin. Struct. Biol.* 21, 189–199.
- Yang, C. W. (2012). A comparative study of short linear motif compositions of the influenza A virus ribonucleoproteins. *PLoS ONE* 7:e38637. doi: 10.1371/journal.pone.0038637
- Zhang, L., Villa, N. Y., Rahman, M. M., Smallwood, S., Shattuck, D., Neff, C., et al. (2009). Analysis of vaccinia virus–host protein–protein interactions: validations of yeast two-hybrid screenings. *J. Proteome Res.* 8, 4311–4318.
- Zhang, Q. C., Petrey, D., Deng, L., Qiang, L., Shi, Y., Thu, C. A., et al. (2012). Structure-based prediction of protein–protein interactions on a genome-wide scale. *Nature* 490, 556–560.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 September 2012; paper pending published: 28 September 2012; accepted: 06 December 2012; published online: 21 December 2012.

Citation: Franzosa EA, Garamszegi S and Xia Y (2012) Toward a three-dimensional view of protein networks between species. *Front. Microbio.* 3:428. doi: 10.3389/fmicb.2012.00428

This article was submitted to *Frontiers in Virology*, a specialty of *Frontiers in Microbiology*.

Copyright © 2012 Franzosa, Garamszegi and Xia. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.