

Northumbria Research Link

Citation: James, Katherine and Munoz, Jose (2022) Computational Network Inference for Bacterial Interactomics. *mSystems*, 7 (2). e01456-21. ISSN 2379-5077

Published by: American Society for Microbiology

URL: <https://doi.org/10.1128/msystems.01456-21>
<<https://doi.org/10.1128/msystems.01456-21>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/48789/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)




**Northumbria
University**
NEWCASTLE



UniversityLibrary



Computational Network Inference for Bacterial Interactomics

 Katherine James,^a Jose Muñoz-Muñoz^a

^aNorthumbria University, Faculty of Health and Life Sciences, Department of Applied Sciences, Newcastle upon Tyne, United Kingdom

ABSTRACT Since the large-scale experimental characterization of protein–protein interactions (PPIs) is not possible for all species, several computational PPI prediction methods have been developed that harness existing data from other species. While PPI network prediction has been extensively used in eukaryotes, microbial network inference has lagged behind. However, bacterial interactomes can be built using the same principles and techniques; in fact, several methods are better suited to bacterial genomes. These predicted networks allow systems-level analyses in species that lack experimental interaction data. This review describes the current network inference and analysis techniques and summarizes the use of computationally-predicted microbial interactomes to date.

KEYWORDS interactome, interologs, data integration, cellular network analysis, systems biology

The representation of biological data as a network, in which nodes represent biological entities and connecting edges represent associations between them, is both visually and computationally tractable (1). Many graph-theoretic tools have been developed to reveal these networks' properties and identify their properties of biological relevance (2–5). Network approaches have been used to investigate maps of protein interactions, termed “interactomes” (6), in several eukaryotic model species using data produced using a number of high-throughput experimental techniques (7–15).

The interactomes of most organisms are largely uncharacterized due to time and cost constraints, and the lack of culture-based technologies for many species. In microbes, the first large-scale bacterial interactome produced was for the gastric pathogen *Helicobacter pylori* using the yeast two-hybrid approach (16), which identifies protein–protein interaction between bait and prey proteins via activation of a reporter gene following reconstitution of its transcription factor. This network was later expanded to cover ~70% of the proteome (17). *Escherichia coli* has by far the most experimental interaction data, including binary protein–protein interaction (PPI) data produced using yeast two-hybrid (18) and protein complex data from affinity purification (19–21). Several large transcriptional, metabolic, and regulatory data sets are also available for this species (22–26). Experimental interaction data sets have been produced for a number of other medically-important pathogenic microbes, including *Campylobacter jejuni* (27), *Mycobacterium tuberculosis* (28), *Mycoplasma pneumoniae* (29), *Streptococcus pneumoniae* (30), and *Treponema pallidum* (31). Interaction data have also been produced for the plant symbiote *Mesorhizobium loti* (32), the model bacterial species *Synechocystis* sp. PCC6803 (33), and *Bacillus subtilis* (34, 35). In addition, a complete complex interactome network has been described for the gut microbiome community describing the microbe–microbe and host–microbe talk elucidating the role of the gut microbiota in health and disease (36).

COMPUTATIONAL INTERACTOME PREDICTION

Several computational PPI prediction methods have been developed that harness existing data (37–39). These methods can be loosely grouped as similarity-based, genome context, evolutionary, and machine learning.

Editor Jack A. Gilbert, University of California San Diego

Copyright © 2022 James and Muñoz-Muñoz. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Katherine James, katherine.p.m.james@northumbria.ac.uk.

The authors declare no conflict of interest.

Accepted 9 March 2022

Published 30 March 2022

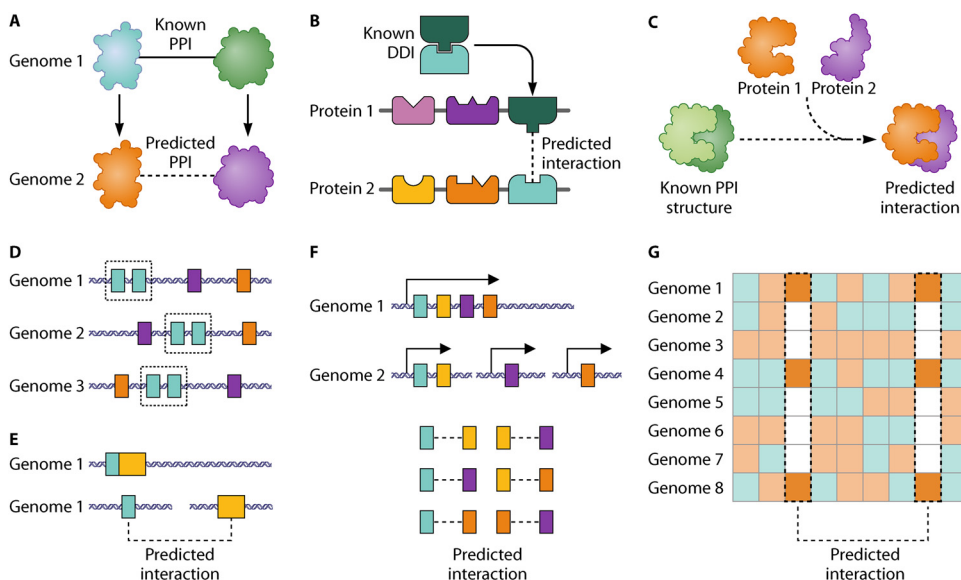


FIG 1 Protein-protein interaction (PPI) inference methods. (A) Interologs: where experimentally confirmed interaction partners in one species have similarity to proteins in another species, an interaction can be predicted. (B) Domain-domain interactions (DDIs): the presence of a pair of domains with a known interaction can be predictive of PPI in other proteins containing those domains. (C) Structural interaction: protein structures can be mapped to the structure of interacting proteins to infer PPI. (D) Gene neighborhood: conservation of protein pairs' (green) proximity in multiple genomes can be predictive of interaction between the pair. (E) Gene fusion: proteins that are fused in one species (yellow and green) have a potential PPI in species in which they are separate proteins. (F) Gene cluster: transcription from an operon in one species indicates functional relation and often PPI in another. Here, an operon of four proteins in one species is predictive of six interactions in another. (G) Phylogenetic profile: protein pairs that interact often have a similar pattern of conservation in multiple genomes (green, presence; orange, absence).

SIMILARITY-BASED

PPIs and their network topology are conserved (40–45). In particular, highly-connected network hub proteins tend to be essential, have slower evolutionary rates, and conserved interactions (3, 46), even between eukaryotes and prokaryotes (47) Protein-protein interactions, termed “interologs,” can consequently be transferred between species (48–52) (Fig. 1A). Conservation has also been observed in regulatory (49, 53), functional (54), and co-expression networks (55–59), allowing transfer of interactions in the same way. Since protein domains are vital to function, the presence of pairs of domains can be predictive of PPI (Fig. 1B), even in proteins with relatively low sequence-similarity (60–62). Domain-domain interaction (DDIs) and interologs are often used in combination to improve network inference (63–67). Similarity of the three-dimensional protein structures can also be used to predict their interactions (Fig. 1C) (68–70). The main prediction methods involve homology modeling (mapping to the known structure of a homologous protein) (71, 72), threading (mapping to known structure of nonhomologous proteins) (73, 74), and docking (predicting the 3D orientation of two interacting proteins) (75–77).

GENOME CONTEXT

Genome context prediction methods compare the location of pairs of genes across multiple genomes (40, 78, 79). “Gene neighborhood” infers interaction based on the assumption that interacting proteins are more closely located on the genome. As an example, in *Bacteroides* spp., genes for SusC-SusD pairs are close in the genome of the bacteria. By comparison of multiple genomes, conservation of pairs' proximity can be identified (Fig. 1D) (80, 81). Although conservation of gene order has been observed in mammals (82), neighborhood-based prediction is most accurate in bacterial genomes (83). “Gene fusion” events can indicate protein interactions since proteins that are fused in one species have a potential functional link in other species in which they are separate proteins (Fig. 1E) (84–86). The “gene cluster” prediction method assumes that

transcription from an operon in one species indicates functional relation and often interaction in another (Fig. 1F) (79, 81).

EVOLUTIONARY

Interacting protein pairs tend to evolve at the same rate (46, 87), and so the distribution of protein pairs will therefore be co-conserved if they interact (78, 88–91). The phylogenetic profile method (Fig. 1G) infers interactions when two genes have a similar pattern of conservation (presence/absence) in multiple genomes (92, 93). MirrorTree and ContextMirror algorithms extend this method beyond binary similarity to globally assess the phylogenies (94, 95). Finally, *in silico* two-hybrid (i2H) method identifies interacting pairs from correlated mutations in multiple sequence alignments, since mutation in interacting proteins co-evolve (96). Several other extensions to these methods have been proposed (97–99).

MACHINE LEARNING

Machine learning can be used to infer PPI by training a classifier on positive (interacting) and negative (noninteracting) pairs of proteins (100). Many data types can be included in the training set including sequence features, co-citation, protein annotation, phylogenetics, expression data, and physiochemical properties (101–109), and multiple data types produce better accuracy than a single input (110, 111). The resulting PPI networks often have confidence scores for putative interactions that allow thresholding and the use of network analysis algorithms that utilize these scores (112). Microbial PPI prediction has been carried out using several algorithms including random forests (113–117), support vector machines (118–120), and Bayesian classifiers (121–123). A related method is probabilistic functional integrated networks (PFINs), which combine multiple data types in a probabilistic framework to produce a network of confidence-weighted interactions (124–126).

NETWORK VALIDATION

Once a predicted interactome has been built, it is essential to evaluate how accurately it represents real cellular biology and assess the level of false interactions that may be present. Evaluation of the quality of a predicted network is difficult due to the level of noise in the underlying data: often several validation methods are required. Small-scale experimental validation can be used on a subset of PPIs to give a level of confidence in the predicted network (20, 64, 122, 127, 128). However, experimental validation is only possible for a small number of interactions. Expression data sets can be used to assess PPIs since interacting proteins are likely to have correlated expression (64, 119, 128–131), and evidence of support for predicted PPIs can be found in other experimental data sets or from text-mined small-scale studies (122, 132–134).

Domain data can be used to assess networks predicted using other input data types (133). Similarly, interaction predictions from other methodologies can be assessed using phylogenetics (119, 122, 129, 131). Hub proteins in predicted networks often correspond to hubs in other species (135). Clustering the network can be used to assess how well the network represents known protein complexes (136). Comparison of the predicted network with random networks can also be used to assess its biological relevance (64, 137).

Protein functional annotations provide useful validation tools as interacting proteins tend to have shared function (64, 119, 128, 131, 132, 138, 139), shared cellular localization (128, 132, 138), and related phenotype (119, 131, 133). Protein functional prediction can provide an objective method of network evaluation by testing its ability to predict the known annotations, for example by a leave-one-out or partitioned cross-validation (140–143). Data partitioning can also be used in the training/testing phase of machine learning to provide a measure of network accuracy (134).

NETWORK ANALYSIS

Several network parameters can be used to reveal aspects of network topology and identify key proteins (144). The *degree* of a protein is its number of interactions; proteins with a high degree are considered hubs and tend to be essential and conserved

(46, 145, 146), and are often targets for pathogens (147, 148). Identification of hubs in predicted networks can highlight important proteins for further study (129). The *degree distribution* of a network, $p(k)$, is the probability a selected protein has k links (149). This distribution reflects the organization of cellular processes (150), with many low degree proteins and a small number of hubs giving a *scale-free* distribution (4, 149, 151). This topology makes networks resistant to random perturbation (5, 151) and has been found in several biological networks in a number of model species (17, 152–154), although some do not have this topology (27, 155).

Biological networks are considered *small world* since they have small diameters (longest shortest path between two proteins) and small characteristic path lengths (average shortest path) relative to equivalently-sized random networks (156–158). The proteins of these networks are arranged in locally-dense regions interconnected by a small number of interactions, and, like scale-free networks, this topology is resistant to perturbation (5).

Several network measures assess the importance of proteins and interactions in networks' information flow to identify bottlenecks. For instance, *betweenness centrality* (159) measures the proportion of shortest paths passing through a protein/interaction. Proteins with high betweenness centrality and low degree often link network modules (160). Betweenness centrality can aid the identification of key proteins within a predicted network (116, 147).

Biological networks tend to have a hierarchical structure of modules within modules (4, 161). Dense network regions are believed to relate to the functional units of the cell (151, 162–165). Partitioning or clustering large networks can reveal the underlying mechanisms of cellular biology and assign protein function (137, 166, 167). Module detection can use additional data, for instance gene expression data (168), functional annotations (169), or domain profiles (170).

Predicted networks can be used to directly annotate proteins with function (171). Network-based annotation transfers known annotations between pairs of directly connected proteins (172, 173), between proteins with shared interaction partners (174), or more globally using network topology (175). Interaction confidence weights, such as those produced by machine learning algorithms, are particularly useful for annotation transfer (173). Annotation and other data can also be used to create process/condition-specific subnetworks (176, 177).

NETWORK COMPARISON

Network comparison can reveal underlying network properties, detect noise, predict missing data, and reveal conserved interactions (41, 178–180). Heuristics, such as global properties and local motifs, are commonly used for comparison (181–183), although some nonheuristics have been developed (184). At the simplest, level biological networks can be compared to network models (185) in which interactions are randomized, while topological characteristics, such as degree distribution and diameter, are preserved, to produce a network profile; similarity of profile indicates underlying similarities of the networks. Networks can be compared directly by comparison of topological properties; however, two networks with similar topology can be vastly different (186). An alternative approach is to analyze the distribution of network motifs (187).

Network alignment produces a more accurate method of comparison (41, 188). Within-species alignment is relatively straightforward since proteins can be merged based on identity (151, 188) and overlap between networks can be used to identify true interactions (189). At a more complex level, networks can be compared across multiple species, either locally by aligning small conserved regions or globally across the network structures (41, 178, 182, 190–194). Alignment complexity increases with the size of the networks and with the number of networks to be aligned (191).

MICROBIAL INTERACTOME NETWORKS

While prokaryotes have far outpaced eukaryotes in the production of sequence data, the opposite is true for interaction data. Largescale experimental data are only

TABLE 1 Predicted bacterial interactomes

Species	Methodology ^a	Proteins	Interactions	Source
<i>Actinobacillus pleuropneumoniae</i>	ORTH	533	2,737	242
<i>Agrobacterium tumefaciens</i>	ORTH	296	690	142
<i>Bacillus anthracis</i>	ORTH	264	732	142
<i>Bacillus licheniformis</i>	ORTH, DDI, GE	2,448	15,864	139
<i>Bacillus subtilis</i>	ORTH	247	707	142
<i>Brucella melitensis</i>	ORTH	238	652	142
<i>Brucella suis</i>	ORTH	225	611	142
<i>Campylobacter jejuni</i>	DDIs	-	-	207
	ORTH	334	1,028	142
<i>Clostridium difficile</i>	RF: STRING, GO	-	955	115
<i>Corynebacterium pseudotuberculosis</i> ^b	ORTH, STRING	-	15,495	200
<i>Escherichia coli</i>	DDIs	-	1,280	207
	ORTH	400	1,473	142
	SVM: GC, CL, PP	3,798	78,122	134
	ML: GC, PP, MT, CM, IH	4,150	1,847,729	121
	EXP, GC	4,146	80,370	20
	PFIN: EXP, DDI, GE, CC, GC, PP	4,099	95,520	124
	PP	1,479	1,618	205
<i>Helicobacter pylori</i>	ORTH	771	5,647	142
<i>Klebsiella pneumoniae</i>	PFIN: ORTH, GC, DDI, PP, GE, CC	4,674	160,450	125
<i>Listeria monocytogenes</i>	ORTH	176	485	142
<i>Methanobrevibacter ruminantium</i>	STRING, GC, MET	637	2,194	202
<i>Methanothermobacter thermautotrophicus</i>	STRING, GC, MET	256	2450	201
<i>Mycobacterium tuberculosis</i>	STRING	3,925	29,664	198
	ORTH	738	5,639	213
	RF: STRING, GO	-	1,854	115
	ORTH, SVM: SEQ	3,465	46,119	119
	STRING	144	587	199
	PP	1,020	911	205
<i>Pseudomonas aeruginosa</i>	ORTH	333	903	142
	RF: GE, CL, GN, DDI, SEQ, FUN	4,181	54,107	116
	PFIN: CC, DDI, GC, GE, ORTH, PP	5,456	203,118	126
<i>Pseudomonas putida</i>	ORTH, DDI	3,254	82,019	210
<i>Salmonella enterica</i> ^b	ORTH, EXP, STRUCT, MET, TF	30,870	81,514	220
<i>Salmonella typhimurium</i>	ORTH	332	1,359	142
<i>Shigella flexneri</i>	ORTH	383	4,548	142
<i>Synechocystis</i> PCC6803	DDI, STRUC, STRING	2,930	109,532	204
	ORTH, DDI, GO	998	8,783	51
	NB: ORTH, DDIs, GC	3,231	4,715	122
<i>Vibrio cholerae</i>	ORTH	275	1,021	142
<i>Vibrio parahaemolyticus</i>	ORTH	365	1,520	142
<i>Vibrio vulnificus</i>	ORTH	372	1,557	142
<i>Yersinia pestis</i>	ORTH	352	1,100	142
<i>Xanthomonas oryzae</i>	ORTH, DDI	1,988	36,886	260

^aCC: co-citation; GE, gene expression; CL, cellular localization; CM, context mirror; DDI, domain-domain interaction; EXP, experimental; FUN, functional interaction; GC, genome context; GO, gene ontology; IH, *in silico* two hybrid; MET, metabolic interactions; ML, machine learning; MT, mirror tree; NB, Naïve Bayes; ORTH, orthology (interologs); PP, phylogenetic profile; PFIN, probabilistic functional integrated network; SEQ, sequence properties; STRING, <https://string-db.org>; STRUC, structural interactions; SVM, support vector machine; TF, transcription factor interactions.

^bCombined for multiple strains.

available in a few species, and small-scale studies require considerable curation to analyze as a whole (195, 196). However, several studies have produced a number of predicted interactomes (Table 1), providing insights into several aspects of microbial biology. Additionally, the STRING database and server contains functional interaction data, including co-citation, co-expression and gene neighborhood, for multiple microbial species (197). STRING data have been used as the initial data source for network studies in several species (115, 178, 198–204).

CELLULAR BIOLOGY AND PROTEIN FUNCTION

In *E. coli*, an interactome study combined experimental data with genome context predictions to assign functions to proteins, including several involved in cell envelope

biogenesis (20). A phylogenetic profile-based network was later produced, which contained previously uncharacterized components of several complexes, including the ribosome (205). Comparison of these networks (20, 205) using edge propagation, demonstrated that both identify complexes overlapping functional modules (206). A support vector machine-derived *E. coli* network was shown to be scale-free and had good overlap with experimental data (134). The EclID database incorporated genome context and phylogenetic evidence into a Bayesian classifier to predict protein function, in particular linking *yeaG* and *yeaH* to nitrogen metabolism (121). EcoliNet is a probabilistic functional integrated network for *E. coli* comprising ~99% of the genome and has successfully predicted knockout phenotypes (124). The interacting domain profile pair method, IDPP, was evaluated on *E. coli* before producing a network for *C. jejuni* from *H. pylori* interaction data (207, 208). IDPP was shown to successfully predict interactions in the target species that were not found in the source species.

Incorporating expression data, interologs and DDIs, *Bacillus licheniformis* proteins were assigned to complexes and putative functions (139). A probabilistic network for *Klebsiella pneumoniae*, derived from multiple data types, was used to identify antibiotic resistance genes (125). Wuchty and colleagues expanded experimental networks of *Streptococcus pneumoniae* and *H. pylori* using interolog data to improve their functional predictive power (30, 209). Interologs and DDIs were combined with experimental data to study the metabolic modules of *Pseudomonas putida* (210). In the related species *P. aeruginosa*, probabilistic functional network integration identified novel virulence and antibiotic resistance genes (126).

SynechoNET is a predicted interactome for *Synechocystis*, focused on membrane biology (204), while InteroPORC predicted an interactome for this species comprising 28% of the genome (51). Later Naïve Bayesian network classification was applied to protein functional prediction and modular analysis in *Synechocystis* (122). Interactome prediction has also been widely used to study *Mycobacterium tuberculosis* due to unavailability of accurate *in vitro* methods in this bacterium (211, 212). Several predicted interactomes have been created to study network properties (119), evolution (205), protein function (115), virulence (213), and drug resistance (198) in this species. Finally, interolog networks for 22 bacterial species (the largest of which are included in Table 1) were produced by McDermott and colleagues and used to predict functions for a large number of unannotated bacterial proteins (142) demonstrating the potential of large-scale network studies to enhance our understanding of bacterial cellular physiology.

Computational prediction has also been used to understand the interplay between bacteriophages and bacteria (214). Phylogenetic profiles using genomic/metagenomic data have identified host–virus and virus–virus relationships (215–218). Leite and colleagues used interactions, DDIs, and sequence properties to compare machine learning frameworks, concluding that predictive power will improve as input data increases (114). More recently, a Markov random field framework of virus–host and virus–virus similarity measures has been developed (219).

INTERACTOME EVOLUTION

Comparison of networks for different species can reveal insights into interactome evolution. Zitnik and colleagues used STRING data to create networks for 1,539 bacteria (178); comparison with those of eukaryotes revealed that interactomes have evolved to become more robust, and that bacterial interactome robustness is associated with more complex environments. Using a binary interaction data set for *Treponema pallidum*, interolog networks were created for 372 other genomes, ~28% of which were estimated to be true interactions (31). This study also revealed that bacterial proteins have higher degrees than eukaryotic proteins. These networks also revealed a central role for cell motility proteins in bacterial interactomes.

By comparing a *Methanobrevibacter ruminantium* network with those of *Methanosarcina acetivorans*, *Methanosarcina barkeri* and *Methanococcus maripaludis* biosynthetic subsystems involved in survival in the rumen were identified (202). By comparing the

Methanothermobacter thermautotrophicus metabolic interactome with those of metal-loving bacteria, separate evolution of niche-specific cellular functions was revealed (201). A comparison of 10 strains of *Salmonella enterica* identified distinct transcription factor targets conferring adaptation to gastrointestinal and extra-intestinal environments (220). Similarly, comparison of host–pathogen interactomes between two strains of *Burkholderia pseudomallei* revealed several interactions unique to the virulent strain and highlighted the potential roles of chaperon and drug/carbohydrate binding proteins during infection (221).

PATHOGEN–HOST INTERACTIONS

Interactome prediction can identify cross talk between pathogen and host (222, 223). A DDI-based network suggested that human–*M. tuberculosis* PPIs tend to have more domains than intraspecies interactions (138), and this trend was later observed in an interolog-based mapping study, which also revealed that hub proteins of intraspecies networks tend to be involved in host–pathogen PPI (148). Using a random forest framework, the cancer pathway was involved in *M. tuberculosis* infection (117), while a DDI network implicated several PPIs involving heat shock, redox proteins (224). Finally, a combination of interolog and DDI mapping associated several genes of the host immune responses to *M. tuberculosis* infection (65).

In *Fusobacterium nucleatum*, a host–pathogen network implicated the Fap2 adhesin as a virulence protein (225). Comparison of machine learning classifiers for *Bacillus anthracis*–humans PPI prediction, suggested neural networks outperform SVMs (120); the resulting interactions revealed involvement of apoptosis and immune regulation pathways in infection. The predicted networks between humans and *B. anthracis*, *Francisella tularensis*, and *Yersinia pestis* indicated that hubs and bottlenecks of the intraspecies networks tend to interact (147, 226). Thirteen membrane proteins of *Leptospira interrogans* were predicted to be involved in cellular disruption during infection, four of which were common between strains (203). Coelho and colleagues produced a human–microbial PPI network of the oral cavity using a Bayesian classifier, which revealed *Rothia mucilaginosa*, *Leptotrichia buccalis*, and *Actinomyces odontolyticus* as having the most interactions with human proteins (123).

In the plant pathogen *Ralstonia solanacearum*, interolog network analysis identified interactions between its transportation proteins and core proteins of the *A. thaliana* interactome (227). The response to metal ions was linked to the host defense response during *Pseudomonas syringae* infection of *A. thaliana* (228). Defense response proteins were also found to be enriched in random forest-derived networks between *A. thaliana* and the pathogens *P. syringae*, *Hyaloperonospora arabidopsis*, and *Golovinomyces orontii* (113). Interolog and DDI mapping has also been used to study plant–bacterial symbiosis, suggesting a role of host 14-3-3 and heat shock proteins in the relationship between *Bradyrhizobium diazoefficiens* and *Glycine max* (229).

VIRUSES

Computational prediction has been applied to viral species including HIV (230–232), hepatitis C virus (233), human papillomaviruses (234), and Ebola (235). In particular, the recent pandemic has highlighted the importance of understanding viral–host interaction, and having resources available to rapidly respond to new viral threats. Two studies have compared PPI prediction to emerging data from Sars-CoV-2: the first used interolog mapping, DDIs, and machine learning to link ACE2 and DPP4 to spike protein binding (236); the second used an ensemble machine learning algorithm based on experimental data and sequence features to predict >1,000 potential human protein targets (237).

DRUG TARGETS

Interactome networks can aid in the identification of potential drug targets by revealing essential pathogen interactions (238). A *M. tuberculosis*–human interactome was shown to be enriched in predicted drug targets (239), and several studies have used computational prediction to identify putative drug targets (199, 213) and to

understand the mechanisms of drug resistance in this important pathogen (198, 240). In *P. aeruginosa* a random forest predicted network was used to prioritize drug targets based on their essentiality and topological importance (116). Using interolog mapping, 12 putative drug targets were identified in methicillin-resistant *Staphylococcus aureus*, including a histone deacetylase (241). A predicted *Corynebacterium pseudotuberculosis* interactome was used to identify 41 essential proteins as candidates for infection diagnosis in livestock, highlighting the tryptophan biosynthesis pathway as a potential drug target (200). The network also revealed that this species may use multiple iron acquisition strategies in low iron environments. In the swine pathogen *Actinobacillus pleuropneumoniae*, nine drug target candidates were identified using interolog mapping (242). Interolog analysis has also been applied to the gut microbiome to identify target species driving metabolic change during disease (243).

CONCLUSIONS AND FUTURE PERSPECTIVES

Although PPI network prediction has been extensively used in eukaryotes, microbial network inference can be achieved using the same principles and analysis techniques. Bacterial interactomes share common hierarchical properties, such as modularity and robustness (244). Many of the caveats to interactome prediction in eukaryotes, such as evolutionary distance, unequal conservation, and physiological context (245–249) are mitigated in prokaryotes due to their smaller genomes and single-celled nature. Several prediction methods, in particular gene neighborhood and gene cluster, are more suited to microbial than eukaryotic use (79, 83), and phylogenetic profiles are powerful predictors, particularly when including inputs from the three domains of life (250).

Interolog and DDI mapping can only detect interactions within conserved areas of the genome (48), and these methods rely on the quality of the underlying interaction data; stochastic activation of reporters can give false positives and low sensitivity leads to false negatives (251, 252), and different methods have their own strengths (253–256). In eukaryotes, poor overlap has been observed between data sets of different types, and between those of the same type (10, 189). Comparison of experimental data of *C. jejuni*, *H. pylori*, and *E. coli* suggests that these data sets have significant levels of overlap and similar rates of false results (27). Meta-interactome analysis can be used to identify broadly-conserved biological systems, although levels of conservation remain low due to lack of experimental interactome coverage in many species (257).

Current experimental interactome data are incomplete and biased toward well-studied proteins and species (178). Using a combination of computational methods (51, 119, 139, 201, 210), and experimental data if available (20, 124, 220), gives a more complete predicted interactome, reduces some biases, and strengthens the evidence of true interactions. Integration of diverse data types is particularly effective when using a probabilistic (124–126) or machine learning (115, 116, 121, 134) framework, allowing thresholding of interaction confidence scores and therefore reduction of noise.

Filling in the gaps in bacterial interactomes is vital to our understanding of their biology, and computational prediction can help to pin down these areas and target further analyses by identifying areas of interest and providing putative protein functions. Microbial comparative interactomics is now possible on a large scale; for instance, the >1,500 predicted networks produced by Zitnik and colleagues revealed the evolutionary rewiring of interactomes through time (178). While there are parts of some interactomes that cannot currently be predicted due to the complexity of the protein locations and the lack of the accurate annotations in the genome of new bacterial isolates, interactome accuracy will improve as coverage of diverse species increases, providing insights in several areas of biology, in particular the identification of PPIs for antibacterial discovery (258), understanding of pathogenicity through host cell rewiring (256), and in engineering of synthetic cellular systems (259).

ACKNOWLEDGMENTS

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Graphic produced by Patrick Lane, ScEYence Studios.

REFERENCES

- Hallinan JS, James K, Wipat A. 2011. Network approaches to the functional analysis of microbial proteins. *Adv Microb Physiol* 59:101–133. <https://doi.org/10.1016/B978-0-12-387661-4.00005-7>.
- Hart G, Lee I, Marcotte ER. 2007. A high-accuracy consensus map of yeast protein complexes reveals modular nature of gene essentiality. *BMC Bioinform* 8:236. <https://doi.org/10.1186/1471-2105-8-236>.
- He X, Zhang J. 2006. Why do hubs tend to be essential in protein networks? *PLoS Genet* 2:e88. <https://doi.org/10.1371/journal.pgen.0020088>.
- Barabási AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5:101–113. <https://doi.org/10.1038/nrg1272>.
- Albert R, Jeong H, Barabási AL. 2000. Error and attack tolerance of complex networks. *Nature* 406:378–382. <https://doi.org/10.1038/35019019>.
- Cusick ME, Klitgord N, Vidal M, Hill DE. 2005. Interactome: gateway into systems biology. *Hum Mol Genet* 14:R171–R181. <https://doi.org/10.1093/hmg/ddi335>.
- Arabidopsis Interactome Mapping Consortium. 2011. Evidence for network evolution in an *Arabidopsis* interactome map. *Science* 333:601–607. <https://doi.org/10.1126/science.1203877>.
- Trigg SA, Garza RM, MacWilliams A, Nery JR, Bartlett A, Castanon R, Goubil A, Feeney J, O'Malley R, Huang S-SC, Zhang ZZ, Galli M, Ecker JR. 2017. CRy2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. *Nat Methods* 14:819–825. <https://doi.org/10.1038/nmeth.4343>.
- Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochar P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadomodar G, Yang M, Johnston M, Fields S, Rothberg JM. 2000. A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae*. *Nature* 403:623–627. <https://doi.org/10.1038/35001009>.
- Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y. 2001. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* 98:4569–4574. <https://doi.org/10.1073/pnas.061034498>.
- Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Hofert C, Schelder M, Brajnovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier M, Copley R, Edelmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415:141–147. <https://doi.org/10.1038/415141a>.
- Gavin A-C, Aloy P, Grandi P, Krause R, Boesche M, Marzioch M, Rau C, Jensen LJ, Bastuck S, Dümpelfeld B, Edelmann A, Heurtier M-A, Hoffman V, Hoefert C, Klein K, Hudak M, Michon A-M, Schelder M, Schirle M, Remor M, Rudi T, Hooper S, Bauer A, Bouwmeester T, Casari G, Drewes G, Neubauer G, Rick JM, Kuster B, Bork P, Russell RB, Superti-Furga G. 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440:631–636. <https://doi.org/10.1038/nature04532>.
- Yu H, Braun P, Yildirim MA, Lemmens I, Venkatesan K, Sahalie J, Hirozane-Kishikawa T, Gebreab F, Li N, Simonis N, Hao T, Rual J-F, Dricot A, Vazquez A, Murray RR, Simon C, Tardivo L, Tam S, Svrikapa N, Fan C, de Smet A-S, Motyl A, Hudson ME, Park J, Xin X, Cusick ME, Moore T, Boone C, Snyder M, Roth FP, Barabási A-L, Tavernier J, Hill DE, Vidal M. 2008. High-quality binary protein interaction map of the yeast interactome network. *Science* 322:104–110. <https://doi.org/10.1126/science.1158684>.
- Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang LV, Wong SL, Franklin G, Li S, Albala JS, Lim J, Fraughton C, Llamasas E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, Smolyar A, Bosak S, Sequerra R, Doucette-Stamm L, Cusick ME, Hill DE, Roth FP, Vidal M. 2005. Towards a proteome-scale map of the human protein–protein interaction network. *Nature* 437:1173–1178. <https://doi.org/10.1038/nature04209>.
- Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, Koeppen S, Timm J, Mintzclaff S, Abraham C, Bock N, Kietzmann S, Goedde A, Toksöz E, Droege A, Krobitsch S, Korn B, Birchmeier W, Lehrach H, Wanker EE. 2005. A human protein–protein interaction network: a resource for annotating the proteome. *Cell* 122:957–968. <https://doi.org/10.1016/j.cell.2005.08.029>.
- Rain JC, Selig L, De Reuse H, Battaglia V, Reverdy C, Simon S, Lenzen G, Petel F, Wojcik J, Schächter V, Chemama Y, Labigne A, Legrain P. 2001. The protein–protein interaction map of *Helicobacter pylori*. *Nature* 409:211–215. <https://doi.org/10.1038/35051615>.
- Häuser R, Ceol A, Rajagopala SV, Mosca R, Siszler G, Wermke N, Sikorski P, Schwarz F, Schick M, Wuchty S, Aloy P, Uetz P. 2014. A second-generation protein–protein interaction network of *Helicobacter pylori*. *Mol Cell Proteomics* 13:1318–1329. <https://doi.org/10.1074/mcp.O113.033571>.
- Rajagopala SV, Sikorski P, Kumar A, Mosca R, Vlasblom J, Arnold R, Franca-Koh J, Pakala SB, Phanse S, Ceol A, Häuser R, Siszler G, Wuchty S, Emili A, Babu M, Aloy P, Pieper R, Uetz P. 2014. The binary protein–protein interaction landscape of *Escherichia coli*. *Nat Biotechnol* 32:285–290. <https://doi.org/10.1038/nbt.2831>.
- Arifuzzaman M, Maeda M, Itoh A, Nishikata K, Takita C, Saito R, Ara T, Nakahigashi K, Huang HC, Hirai A, Tsuzuki K, Nakamura S, Altaf-Ul-Amin M, Oshima T, Baba T, Yamamoto N, Kawamura T, Ioka-Nakamichi T, Kitagawa M, Tomita M, Kanaya S, Wada C, Mori H. 2006. Large-scale identification of protein–protein interaction of *Escherichia coli* K-12. *Genome Res* 16:686–691. <https://doi.org/10.1101/gr.4527806>.
- Hu P, Janga SC, Babu M, Díaz-Mejía JJ, Butland G, Yang W, Pogoutse O, Guo X, Phanse S, Wong P, Chandran S, Christopoulos C, Nazarians-Armavil A, Nasser NK, Musso G, Ali M, Nazemof N, Eroukova V, Golshani A, Paccanaro A, Greenblatt JF, Moreno-Hagelsieb G, Emili A. 2009. Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol* 7:e96. <https://doi.org/10.1371/journal.pbio.1000096>.
- Babu M, Bundalovic-Torma C, Calmettes C, Phanse S, Zhang Q, Jiang Y, Minic Z, Kim S, Mehla J, Gagarianova A, Rodionova I, Kumar A, Guo H, Kagan O, Pogoutse O, Aoki H, Deineko V, Caufield JH, Holtzapfel E, Zhang Z, Vastermark A, Pandya Y, Lai CC-L, El Bakkouri M, Hooda Y, Shah M, Burnside D, Hooshyar M, Vlasblom J, Rajagopala SV, Golshani A, Wuchty S, F Greenblatt J, Saier M, Uetz P, F Moraes T, Parkinson J, Emili A. 2018. Global landscape of cell envelope protein complexes in *Escherichia coli*. *Nat Biotechnol* 36:103–112. <https://doi.org/10.1038/nbt.4024>.
- Resendis-Antonio O, Freyre-González JA, Menchaca-Méndez R, Gutiérrez-Rios RM, Martínez-Antonio A, Avila-Sánchez C, Collado-Vides J. 2005. Modular analysis of the transcriptional regulatory network of *E. coli*. *Trends Genet* 21:16–20. <https://doi.org/10.1016/j.tig.2004.11.010>.
- Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, Kasif S, Collins JJ, Gardner TS. 2007. Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol* 5:e8. <https://doi.org/10.1371/journal.pbio.0050008>.
- Barrett CL, Herring CD, Reed JL, Palsson BO. 2005. The global transcriptional regulatory network for metabolism in *Escherichia coli* exhibits few dominant functional states. *Proc Natl Acad Sci U S A* 102:19103–19108. <https://doi.org/10.1073/pnas.0505231102>.
- Fang X, Sastry A, Mih N, Kim D, Tan J, Yurkovich JT, Lloyd CJ, Gao Y, Yang L, Palsson BO. 2017. Global transcriptional regulatory network for *Escherichia coli* robustly connects gene expression to transcription factor activities. *Proc Natl Acad Sci U S A* 114:10286–10291. <https://doi.org/10.1073/pnas.1702581114>.
- Lempp M, Farke N, Kuntz M, Freibert SA, Lill R, Link H. 2019. Systematic identification of metabolites controlling gene expression in *E. coli*. *Nat Commun* 10:4463. [Mismatch] <https://doi.org/10.1038/s41467-019-12474-1>.
- Parrish JR, Yu J, Liu G, Hines JA, Chan JE, Mangiola BA, Zhang H, Pacifico S, Fotouhi F, DiRita VJ, Ideker T, Andrews P, Finley RL. 2007. A proteome-wide protein interaction map for *Campylobacter jejuni*. *Genome Biol* 8:R130. <https://doi.org/10.1186/gb-2007-8-7-r130>.
- Wang Y, Cui T, Zhang C, Yang M, Huang Y, Li W, Zhang L, Gao C, He Y, Li Y, Huang F, Zeng J, Huang C, Yang Q, Tian Y, Zhao C, Chen H, Zhang H, He ZG. 2010. Global protein–protein interaction network in the human pathogen *Mycobacterium tuberculosis* H37Rv. *J Proteome Res* 9:6665–6677. <https://doi.org/10.1021/pr100808n>.

29. Kühner S, van Noort V, Betts MJ, Leo-Macias A, Batisse C, Rode M, Yamada T, Maier T, Bader S, Beltran-Alvarez P, Castaño-Diez D, Chen WH, Devos D, Güell M, Norambuena T, Racke I, Rybin V, Schmidt A, Yus E, Aebbers R, Herrmann R, Böttcher B, Frangakis AS, Russell RB, Serrano L, Bork P, Gavin AC. 2009. Proteome organization in a genome-reduced bacterium. *Science* 326:1235–1240. <https://doi.org/10.1126/science.1176343>.
30. Wuchty S, Rajagopala SV, Blazie SM, Parrish JR, Khuri S, Finley RL, Uetz P. 2017. The protein interactome of *Streptococcus pneumoniae* and bacterial meta-interactomes improve function predictions. *mSystems* 2: e00019–17. <https://doi.org/10.1128/mSystems.00019-17>.
31. Titz B, Rajagopala SV, Goll J, Häuser R, McKeivitt MT, Palzkill T, Uetz P. 2008. The binary protein interactome of *Treponema pallidum*—the syphilis spirochete. *PLoS One* 3:e2292. <https://doi.org/10.1371/journal.pone.0002292>.
32. Shimoda Y, Shinpo S, Kohara M, Nakamura Y, Tabata S, Sato S. 2008. A large scale analysis of protein–protein interactions in the nitrogen-fixing bacterium *Mesorhizobium loti*. *DNA Res* 15:13–23. <https://doi.org/10.1093/dnares/dsm028>.
33. Sato S, Shimoda Y, Muraki A, Kohara M, Nakamura Y, Tabata S. 2007. A large-scale protein–protein interaction analysis in *Synechocystis* sp. PCC6803. *DNA Res* 14:207–216. <https://doi.org/10.1093/dnares/dsm021>.
34. Marchadier E, Carballido-López R, Brinster S, Fabret C, Mervelet P, Bessières P, Noirot-Gros M-F, Fromion V, Noirot P. 2011. An expanded protein–protein interaction network in *Bacillus subtilis* reveals a group of hubs: exploration by an integrative approach. *Proteomics* 11:2981–2991. <https://doi.org/10.1002/pmic.201000791>.
35. Noirot-Gros MF, Dervyn E, Wu LJ, Mervelet P, Errington J, Ehrlich SD, Noirot P. 2002. An expanded view of bacterial DNA replication. *Proc Natl Acad Sci U S A* 99:8342–8347. <https://doi.org/10.1073/pnas.122040799>.
36. Sung J, Kim S, Cabatbat JJT, Jang S, Jin YS, Jung GY, Chia N, Kim PJ. 2017. Global metabolic interaction network of the human gut microbiota for context-specific community-scale analysis. *Nat Comm* 8:15393. <https://doi.org/10.1038/ncomms15393>.
37. Skrabanek L, Saini HK, Bader GD, Enright AJ. 2008. Computational prediction of protein–protein interactions. *Mol Biotechnol* 38:1–17. <https://doi.org/10.1007/s12033-007-0069-2>.
38. Liu ZP, Chen L. 2012. Proteome-wide prediction of protein–protein interactions from high-throughput data. *Protein Cell* 3:508–520. <https://doi.org/10.1007/s13238-012-2945-1>.
39. Khatun MS, Shoombatong W, Hasan MM, Kurata H. 2020. Evolution of sequence-based bioinformatics tools for protein–protein interaction prediction. *Curr Genomics* 21:454–463. <https://doi.org/10.2174/1389202921999200625103936>.
40. von Mering C, Zdobnov EM, Tsoka S, Ciccarelli FD, Pereira-Leal JB, Ouzounis CA, Bork P. 2003. Genome evolution reveals biochemical networks and functional modules. *Proc Natl Acad Sci U S A* 100: 15428–15433. <https://doi.org/10.1073/pnas.2136809100>.
41. Sharan R, Suthram S, Kelley RM, Kuhn T, McCuine S, Uetz P, Sittler T, Karp RM, Ideker T. 2005. Conserved patterns of protein interaction in multiple species. *Proc Natl Acad Sci U S A* 102:1974–1979. <https://doi.org/10.1073/pnas.0409522102>.
42. Boltz TA, Devkota P, Wuchty S. 2019. Collective influencers in protein interaction networks. *Sci Rep* 9:3948. <https://doi.org/10.1038/s41598-019-40410-2>.
43. Wuchty S, Barabási AL, Ferdig MT. 2006. Stable evolutionary signal in a yeast protein interaction network. *BMC Evol Biol* 6:8. <https://doi.org/10.1186/1471-2148-6-8>.
44. Hirsh E, Sharan R. 2007. Identification of conserved protein complexes based on a model of protein network evolution. *Bioinformatics* 23: e170–e176. <https://doi.org/10.1093/bioinformatics/btl295>.
45. Mahajan T, Dar RD. 2021. Internetwork connectivity of molecular networks across species of life. *Sci Rep* 11:1168. <https://doi.org/10.1038/s41598-020-80745-9>.
46. Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW. 2002. Evolutionary rate in the protein interaction network. *Science* 296:750–752. <https://doi.org/10.1126/science.1068696>.
47. Kelley BP, Sharan R, Karp RM, Sittler T, Root DE, Stockwell BR, Ideker T. 2003. Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proc Natl Acad Sci U S A* 100: 11394–11399. <https://doi.org/10.1073/pnas.1534710100>.
48. Matthews LR, Vaglio P, Reboul J, Ge H, Davis BP, Garrels J, Vincent S, Vidal M. 2001. Identification of potential interaction networks using sequence-based searches for conserved protein–protein interactions or “interologs.” *Genome Res* 11:2120–2126. <https://doi.org/10.1101/gr.205301>.
49. Yu H, Luscombe NM, Lu HX, Zhu X, Xia Y, Han JDJ, Bertin N, Chung S, Vidal M, Gerstein M. 2004. Annotation transfer between genomes: protein–protein interologs and protein–DNA regulogs. *Genome Res* 14: 1107–1118. <https://doi.org/10.1101/gr.1774904>.
50. Lee SA, Chan C, Tsai CH, Lai JM, Wang FS, Kao CY, Huang CYF. 2008. Ortholog-based protein–protein interaction prediction and its application to inter-species interactions. *BMC Bioinform* 9:S11. <https://doi.org/10.1186/1471-2105-9-S12-S11>.
51. Michaut M, Kerrien S, Montecchi-Palazzi L, Chauvat F, Cassier-Chauvat C, Aude J-C, Legrain P, Hermjakob H. 2008. InteroPORC: automated inference of highly conserved protein interaction networks. *Bioinformatics* 24:1625–1631. <https://doi.org/10.1093/bioinformatics/btn249>.
52. Wiles AM, Doderer M, Ruan J, Gu TT, Ravi D, Blackman B, Bishop AJR. 2010. Building and analyzing protein interactome networks by cross-species comparisons. *BMC Syst Biol* 4:36. <https://doi.org/10.1186/1752-0509-4-36>.
53. Baumbach J. 2010. On the power and limits of evolutionary conservation-unraveling bacterial gene regulatory networks. *Nucleic Acids Res* 38:7877–7884. <https://doi.org/10.1093/nar/gkq699>.
54. Kim E, Bae D, Yang S, Ko G, Lee S, Lee B, Lee I. 2020. BiomeNet: a database for construction and analysis of functional interaction networks for any species with a sequenced genome. *Bioinformatics* 36:1584–1589. <https://doi.org/10.1093/bioinformatics/btz776>.
55. Tsaparas P, Mariño-Ramírez L, Bodenreider O, Koonin EV, Jordan IK. 2006. Global similarity and local divergence in human and mouse gene co-expression networks. *BMC Evol Biol* 6:70. <https://doi.org/10.1186/1471-2148-6-70>.
56. Stuart JM, Segal E, Koller D, Kim SK. 2003. A gene-coexpression network for global discovery of conserved genetic modules. *Science* 302: 249–255. <https://doi.org/10.1126/science.1087447>.
57. van Noort V, Snel B, Huynen MA. 2003. Predicting gene function by conserved co-expression. *Trends Genet* 19:238–242. [https://doi.org/10.1016/S0168-9525\(03\)00056-8](https://doi.org/10.1016/S0168-9525(03)00056-8).
58. Snel B, van Noort V, Huynen MA. 2004. Gene co-regulation is highly conserved in the evolution of eukaryotes and prokaryotes. *Nucleic Acids Res* 32:4725–4731. <https://doi.org/10.1093/nar/gkh815>.
59. Alkema WBL, Lenhard B, Wasserman WW. 2004. Regulog analysis: detection of conserved regulatory networks across bacteria: application to *Staphylococcus aureus*. *Genome Res* 14:1362–1373. <https://doi.org/10.1101/gr.2242604>.
60. Sprinzak E, Margalit H. 2001. Correlated sequence-signatures as markers of protein–protein interaction. *J Mol Biol* 311:681–692. <https://doi.org/10.1006/jmbi.2001.4920>.
61. Deng M, Mehta S, Sun F, Chen T. 2002. Inferring domain–domain interactions from protein–protein interactions. *Genome Res* 12:1540–1548. <https://doi.org/10.1101/gr.153002>.
62. Ng SK, Zhang Z, Tan SH. 2003. Integrative approach for computationally inferring protein domain interactions. *Bioinformatics* 19:923–929. <https://doi.org/10.1093/bioinformatics/btg118>.
63. Zhao XM, Zhang XW, Tang WH, Chen L. 2009. FPPi: *Fusarium graminearum* protein–protein interaction database. *J Proteome Res* 8:4714–4721. <https://doi.org/10.1021/pr900415b>.
64. Zhang K, Li Y, Li T, Li ZG, Hsiang T, Zhang Z, Sun W. 2017. Pathogenicity genes in *Ustilagoideia virens* revealed by a predicted protein–protein interaction network. *J Proteome Res* 16:1193–1206. <https://doi.org/10.1021/acs.jproteome.6b00720>.
65. Huo T, Liu W, Guo Y, Yang C, Lin J, Rao Z. 2015. Prediction of host–pathogen protein interactions between *Mycobacterium tuberculosis* and *Homo sapiens* using sequence motifs. *BMC Bioinform* 16:100. <https://doi.org/10.1186/s12859-015-0535-y>.
66. Lei D, Lin R, Yin C, Li P, Zheng A. 2014. Global protein–protein interaction network of rice sheath blight pathogen. *J Proteome Res* 13:3277–3293. <https://doi.org/10.1021/pr500069r>.
67. Singh G, Singh V, Singh V. 2021. Genome-wide interologous interactome map (TeaGPIN) of *Camellia sinensis*. *Genomics* 113:553–564. <https://doi.org/10.1016/j.ygeno.2020.09.048>.
68. Lo YS, Chen YC, Yang JM. 2010. 3D-interologs: an evolution database of physical protein–protein interactions across multiple genomes. *BMC Genome* 11:S7. <https://doi.org/10.1186/1471-2164-11-S3-S7>.
69. Dror O, Schneidman-Duhovny D, Shulman-Peleg A, Nussinov R, Wolfson HJ, Sharan R. 2008. Structural similarity of genetically interacting proteins. *BMC Syst Biol* 2:69. <https://doi.org/10.1186/1752-0509-2-69>.

70. Ogmen U, Keskin O, Aytuna AS, Nussinov R, Gursoy A. 2005. PRISM: protein interactions by structural matching. *Nucleic Acids Res* 33: W331–W336. <https://doi.org/10.1093/nar/gki585>.
71. Aloy P, Russell RB. 2003. InterPreTS: protein interaction prediction through tertiary structure. *Bioinformatics* 19:161–162. <https://doi.org/10.1093/bioinformatics/19.1.161>.
72. Aloy P, Böttcher B, Ceulemans H, Leutwein C, Mellwig C, Fischer S, Gavin AC, Bork P, Superti-Furga G, Serrano L, Russell RB. 2004. Structure-based assembly of protein complexes in yeast. *Science* 303:2026–2029. <https://doi.org/10.1126/science.1092645>.
73. Lu L, Arakaki AK, Lu H, Skolnick J. 2003. Multimeric threading-based prediction of protein–protein interactions on a genomic scale: application to the *Saccharomyces cerevisiae* proteome. *Genome Res* 13:1146–1154. <https://doi.org/10.1101/gr.1145203>.
74. Lu L, Lu H, Skolnick J. 2002. MULTIPROSPECTOR: an algorithm for the prediction of protein–protein interactions by multimeric threading. *Proteins* 49:350–364. <https://doi.org/10.1002/prot.10222>.
75. Janin J, Henrick K, Moult J, Eyck LT, Sternberg MJ, Vajda S, Vakser I, Wodak SJ. 2003. CAPRI: a Critical Assessment of PRedicted Interactions. *Proteins* 52:2–9. <https://doi.org/10.1002/prot.10381>.
76. Smith GR, Sternberg MJ. 2002. Prediction of protein–protein interactions by docking methods. *Curr Opin Struct Biol* 12:28–35. [https://doi.org/10.1016/s0959-440x\(02\)00285-3](https://doi.org/10.1016/s0959-440x(02)00285-3).
77. Ohue M, Matsuzaki Y, Shimoda T, Ishida T, Akiyama Y. 2013. Highly precise protein–protein interaction prediction based on consensus between template-based and *de novo* docking methods. *BMC Proc* 7:56. <https://doi.org/10.1186/1753-6561-7-57-S6>.
78. Sun J, Sun Y, Ding G, Liu Q, Wang C, He Y, Shi T, Li Y, Zhao Z. 2007. InPrePPI: an integrated evaluation method based on genomic context for predicting protein–protein interactions in prokaryotic genomes. *BMC Bioinform* 8:414. <https://doi.org/10.1186/1471-2105-8-414>.
79. Bowers PM, Pellegrini M, Thompson MJ, Fierro J, Yeates TO, Eisenberg D. 2004. Prolinks: a database of protein functional linkages derived from coevolution. *Genome Biol* 5:R35. <https://doi.org/10.1186/gb-2004-5-5-r35>.
80. Dandekar T, Snel B, Huynen M, Bork P. 1998. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci* 23: 324–328. [https://doi.org/10.1016/s0968-0004\(98\)01274-2](https://doi.org/10.1016/s0968-0004(98)01274-2).
81. Overbeek R, Fonstein M, D'Souza M, Pusch GD, Maltsev N. 1999. The use of gene clusters to infer functional coupling. *Proc Natl Acad Sci U S A* 96: 2896–2901. <https://doi.org/10.1073/pnas.96.6.2896>.
82. Kirk IK, Weinhold N, Brunak S, Belling K. 2017. The impact of the protein interactome on the syntenic structure of mammalian genomes. *PLoS One* 12:e0179112. <https://doi.org/10.1371/journal.pone.0179112>.
83. Huynen M, Snel B, Lathe W, Bork P. 2000. Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. *Genome Res* 10:1204–1210. <https://doi.org/10.1101/gr.10.8.1204>.
84. Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D. 1999. Detecting protein function and protein–protein interactions from genome sequences. *Science* 285:751–753. <https://doi.org/10.1126/science.285.5428.751>.
85. Enright AJ, Iliopoulos I, Kyripides NC, Ouzounis CA. 1999. Protein interaction maps for complete genomes based on gene fusion events. *Nature* 402:86–90. <https://doi.org/10.1038/47056>.
86. Yanai I, Derti A, DeLisi C. 2001. Genes linked by fusion events are generally of the same functional category: a systematic analysis of 30 microbial genomes. *Proc Natl Acad Sci U S A* 98:7940–7945. <https://doi.org/10.1073/pnas.141236298>.
87. Goh CS, Bogan AA, Joachimiak M, Walther D, Cohen FE. 2000. Co-evolution of proteins with their interaction partners. *J Mol Biol* 299:283–293. <https://doi.org/10.1006/jmbi.2000.3732>.
88. Goh CS, Cohen FE. 2002. Co-evolutionary analysis reveals insights into protein–protein interactions. *J Mol Biol* 324:177–192. [https://doi.org/10.1016/s0022-2836\(02\)01038-0](https://doi.org/10.1016/s0022-2836(02)01038-0).
89. Pazos F, Ranea JA, Juan D, Sternberg MJ. 2005. Assessing protein co-evolution in the context of the tree of life assists in the prediction of the interactome. *J Mol Biol* 352:1002–1015. <https://doi.org/10.1016/j.jmb.2005.07.005>.
90. Wu J, Kasif S, DeLisi C. 2003. Identification of functional links between genes using phylogenetic profiles. *Bioinformatics* 19:1524–1530. <https://doi.org/10.1093/bioinformatics/btg187>.
91. Jothi R, Kann MG, Przytycka TM. 2005. Predicting protein–protein interaction by searching evolutionary tree automorphism space. *Bioinformatics* 21:i241–i250. <https://doi.org/10.1093/bioinformatics/bti1009>.
92. Gaasterland T, Ragan MA. 1998. Microbial genomes: phyletic and functional patterns of ORF distribution among prokaryotes. *Microb Comp Genomics* 3:199–217. <https://doi.org/10.1089/omi.1.1998.3.199>.
93. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO. 1999. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc Natl Acad Sci U S A* 96:4285–4288. <https://doi.org/10.1073/pnas.96.8.4285>.
94. Pazos F, Valencia A. 2001. Similarity of phylogenetic trees as indicator of protein–protein interaction. *Protein Eng* 14:609–614. <https://doi.org/10.1093/protein/14.9.609>.
95. Juan D, Pazos F, Valencia A. 2008. High-confidence prediction of global interactomes based on genome-wide coevolutionary networks. *Proc Natl Acad Sci U S A* 105:934–939. <https://doi.org/10.1073/pnas.0709671105>.
96. Pazos F, Valencia A. 2002. In silico two-hybrid system for the selection of physically interacting protein pairs. *Proteins* 47:219–227. <https://doi.org/10.1002/prot.10074>.
97. de Vienne DM, Azé J. 2012. Efficient prediction of co-complexed proteins based on coevolution. *PLoS One* 7:e48728. <https://doi.org/10.1371/journal.pone.0048728>.
98. Date SV, Marcotte EM. 2003. Discovery of uncharacterized cellular systems by genome-wide analysis of functional linkages. *Nat Biotechnol* 21: 1055–1062. <https://doi.org/10.1038/nbt861>.
99. Sato T, Yamanishi Y, Kanehisa M, Toh H. 2005. The inference of protein–protein interactions by co-evolutionary analysis is improved by excluding the information about the phylogenetic relationships. *Bioinformatics* 21:3482–3489. <https://doi.org/10.1093/bioinformatics/bti564>.
100. Qi Y, Bar-Joseph Z, Klein-Seetharaman J. 2006. Evaluation of different biological data and computational classification methods for use in protein interaction prediction. *Proteins* 63:490–500. <https://doi.org/10.1002/prot.20865>.
101. Yamanishi Y, Vert JP, Kanehisa M. 2004. Protein network inference from multiple genomic data: a supervised approach. *Bioinformatics* 20: i363–i370. <https://doi.org/10.1093/bioinformatics/bth910>.
102. Scott MS, Barton GJ. 2007. Probabilistic prediction and ranking of human protein–protein interactions. *BMC Bioinform* 8:239. <https://doi.org/10.1186/1471-2105-8-239>.
103. García-Jiménez B, Juan D, Ezkurdia I, Andrés-León E, Valencia A. 2010. Inference of functional relations in predicted protein networks with a machine learning approach. *PLoS One* 5:e9969. <https://doi.org/10.1371/journal.pone.0009969>.
104. Ben-Hur A, Noble WS. 2005. Kernel methods for predicting protein–protein interactions. *Bioinformatics* 21:i38–i46. <https://doi.org/10.1093/bioinformatics/bti1016>.
105. Nguyen TP, Ho TB. 2008. An integrative domain-based approach to predicting protein–protein interactions. *J Bioinform Comput Biol* 6: 1115–1132. <https://doi.org/10.1142/s0219720008003874>.
106. Hue M, Riffle M, Vert JP, Noble WS. 2010. Large-scale prediction of protein–protein interactions from structures. *BMC Bioinform* 11:144. <https://doi.org/10.1186/1471-2105-11-144>.
107. Chen XW, Liu M. 2005. Prediction of protein–protein interactions using random decision forest framework. *Bioinformatics* 21:4394–4400. <https://doi.org/10.1093/bioinformatics/bti721>.
108. Murakami Y, Mizuguchi K. 2014. Homology-based prediction of interactions between proteins using averaged one-dependence estimators. *BMC Bioinform* 15:213. <https://doi.org/10.1186/1471-2105-15-213>.
109. Chen KH, Wang TF, Hu YJ. 2019. Protein–protein interaction prediction using a hybrid feature representation and a stacked generalization scheme. *BMC Bioinform* 20:308. <https://doi.org/10.1186/s12859-019-2907-1>.
110. Aho T, Almusá H, Matilainen J, Larjo A, Ruusuvoori P, Aho KL, Wilhelm T, Lähdesmäki H, Beyer A, Harju M, Chowdhury S, Leinonen K, Roos C, Yli-Harja O. 2010. Reconstruction and validation of RefRec: a global model for the yeast molecular interaction network. *PLoS One* 5:e10662. <https://doi.org/10.1371/journal.pone.0010662>.
111. Liu Y, Kim I, Zhao H. 2008. Protein interaction predictions from diverse sources. *Drug Discov Today* 13:409–416. <https://doi.org/10.1016/j.drudis.2008.01.005>.
112. Rhodes DR, Tomlins SA, Varambally S, Mahavisno V, Barrette T, Kalyana-Sundaram S, Ghosh D, Pandey A, Chinnaiyan AM. 2005. Probabilistic model of the human protein–protein interaction network. *Nat Biotechnol* 23:951–959. <https://doi.org/10.1038/nbt1103>.
113. Yang S, Li H, He H, Zhou Y, Zhang Z. 2019. Critical assessment and performance improvement of plant-pathogen protein–protein interaction prediction methods. *Brief Bioinform* 20:274–287. <https://doi.org/10.1093/bib/bbx123>.

114. Leite DMC, Brochet X, Resch G, Que YA, Neves A, Peña-Reyes C. 2018. Computational prediction of inter-species relationships through omics data analysis and machine learning. *BMC Bioinform* 19:420. <https://doi.org/10.1186/s12859-018-2388-7>.
115. Ananthasubramanian S, Metri R, Khetan A, Gupta A, Handen A, Chandra N, Ganapathiraju M. 2012. *Mycobacterium tuberculosis* and *Clostridium difficile* interactomes: demonstration of rapid development of computational system for bacterial interactome prediction. *Microb Infect Exp* 2:4. <https://doi.org/10.1186/2042-5783-2-4>.
116. Zhang M, Su S, Bhatnagar RK, Hassett DJ, Lu LJ. 2012. Prediction and analysis of the protein interactome in *Pseudomonas aeruginosa* to enable network-based drug target selection. *PLoS One* 7:e41202. <https://doi.org/10.1371/journal.pone.0041202>.
117. Sun J, Yang LL, Chen X, Kong DX, Liu R. 2018. Integrating multifaceted information to predict *Mycobacterium tuberculosis*–human protein–protein interactions. *J Proteome Res* 17:3810–3823. <https://doi.org/10.1021/acs.jproteome.8b00497>.
118. Bock JR, Gough DA. 2001. Predicting protein–protein interactions from primary structure. *Bioinformatics* 17:455–460. <https://doi.org/10.1093/bioinformatics/17.5.455>.
119. Liu ZP, Wang J, Qiu YQ, Leung RKK, Zhang XS, Tsui SKW, Chen L. 2012. Inferring a protein interaction map of *Mycobacterium tuberculosis* based on sequences and interologs. *BMC Bioinform* 13:56. <https://doi.org/10.1186/1471-2105-13-56>.
120. Ahmed I, Witbooi P, Christoffels A. 2018. Prediction of human–*Bacillus anthracis* protein–protein interactions using multi-layer neural network. *Bioinformatics* 34:4159–4164. <https://doi.org/10.1093/bioinformatics/bty504>.
121. Andres Leon E, Ezkurdia I, Garcia B, Valencia A, Juan D. 2009. ECLID: a database for the inference of functional interactions in *E. coli*. *Nucleic Acids Res* 37:D629–D635. <https://doi.org/10.1093/nar/gkn853>.
122. Lv Q, Ma W, Liu H, Li J, Wang H, Lu F, Zhao C, Shi T. 2015. Genome-wide protein–protein interactions and protein function exploration in cyanobacteria. *Sci Rep* 5:15519. <https://doi.org/10.1038/srep15519>.
123. Coelho ED, Arrais JP, Matos S, Pereira C, Rosa N, Correia MJ, Barros M, Oliveira JL. 2014. Computational prediction of the human-microbial oral interactome. *BMC Syst Biol* 8:24. <https://doi.org/10.1186/1752-0509-8-24>.
124. Kim H, Shim JE, Shin J, Lee I. 2015. EcoliNet: a database of cofunctional gene network for *Escherichia coli*. *Database (Oxford)* 2015:bav001. <https://doi.org/10.1093/database/bav001>.
125. Lee M, Pinto NA, Kim CY, Yang S, D'Souza R, Yong D, Lee I. 2019. Network integrative genomic and transcriptomic analysis of carbapenem-resistant *Klebsiella pneumoniae* strains identifies genes for antibiotic resistance and virulence. *mSystems* 4:2379–5077. <https://doi.org/10.1128/mSystems.00202-19>.
126. Hwang S, Kim CY, Ji SG, Go J, Kim H, Yang S, Kim HJ, Cho A, Yoon SS, Lee I. 2016. Network-assisted investigation of virulence and antibiotic-resistance systems in *Pseudomonas aeruginosa*. *Sci Rep* 6:26223. <https://doi.org/10.1038/srep26223>.
127. Liu X, Huang Y, Liang J, Zhang S, Li Y, Wang J, Shen Y, Xu Z, Zhao Y. 2014. Computational prediction of protein interactions related to the invasion of erythrocytes by malarial parasites. *BMC Bioinform* 15:393. <https://doi.org/10.1186/s12859-014-0393-z>.
128. Wang TY, He F, Hu QW, Zhang Z. 2011. A predicted protein–protein interaction network of the filamentous fungus *Neurospora crassa*. *Mol Biosyst* 7:2278–2285. <https://doi.org/10.1039/c1mb05028a>.
129. Liu ZP, Wang J, Qiu YQ, Leung RKK, Zhang XS, Tsui SKW, Chen L. 2012. Inferring protein–protein interactions based on sequences and interologs in *Mycobacterium tuberculosis*, p 91–96. In Huang D-S, Gan Y, Premaratne P, Han K (ed), *Bio-inspired computing and applications*. Springer, Berlin, Germany.
130. Jansen R, Greenbaum D, Gerstein M. 2002. Relating whole-genome expression data with protein–protein interactions. *Genome Res* 12:37–46. <https://doi.org/10.1101/gr.205602>.
131. Alkan F, Erten C. 2015. SiPAN: simultaneous prediction and alignment of protein–protein interaction networks. *Bioinformatics* 31:2356–2363. <https://doi.org/10.1093/bioinformatics/btv160>.
132. Urban M, Lysenko A, Janowska-Sedja EI, Rawlings C, Hammond-Kosack KE, Tsoka S. 2019. PHI-Nets: a network resource for Ascomycete fungal pathogens to annotate and identify putative virulence interacting proteins and siRNAs. *Front Microbiol* 10:2721. <https://doi.org/10.3389/fmicb.2019.02721>.
133. Liu S, Liu Y, Zhao J, Cai S, Qian H, Zuo K, Zhao L, Zhang L. 2017. A computational interactome for prioritizing genes associated with complex agronomic traits in rice (*Oryza sativa*). *Plant J* 90:177–188. <https://doi.org/10.1111/tpj.13475>.
134. Yellaboina S, Goyal K, Mande SC. 2007. Inferring genome-wide functional linkages in *E. coli* by combining improved genome context methods: comparison with high-throughput experimental data. *Genome Res* 17:527–535. <https://doi.org/10.1101/gr.5900607>.
135. Wuchty S. 2007. Rich-club phenomenon in the interactome of *P. falciparum*—artifact or signature of a parasitic life style? *PLoS One* 2:e335. <https://doi.org/10.1371/journal.pone.0000335>.
136. LaCount DJ, Vignali M, Chettier R, Phansalkar A, Bell R, Hesselberth JR, Schoenfeld LW, Ota I, Sahasrabudhe S, Kurschner C, Fields S, Hughes RE. 2005. A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* 438:103–107. <https://doi.org/10.1038/nature04104>.
137. Brohée S, van Helden J. 2006. Evaluation of clustering algorithms for protein–protein interaction networks. *BMC Bioinform* 7:488. <https://doi.org/10.1186/1471-2105-7-488>.
138. Zhou H, Rezaei J, Hugo W, Gao S, Jin J, Fan M, Yong CH, Wozniak M, Wong L. 2013. Stringent DDI-based prediction of *H. sapiens*–*M. tuberculosis* H37Rv protein–protein interactions. *BMC Syst Biol* 7:56. <https://doi.org/10.1186/1752-0509-7-56-56>.
139. Han YC, Song JM, Wang L, Shu CC, Guo J, Chen LL. 2016. Prediction and characterization of protein–protein interaction network in *Bacillus licheniformis* WX-02. *Sci Rep* 6:19486. <https://doi.org/10.1038/srep19486>.
140. Deng M, Tu Z, Sun F, Chen T. 2004. Mapping gene ontology to proteins based on protein–protein interaction data. *Bioinformatics* 20:895–902. <https://doi.org/10.1093/bioinformatics/btg500>.
141. Nariai N, Kolaczyk ED, Kasif S. 2007. Probabilistic protein function prediction from heterogeneous genome-wide data. *PLoS One* 2:e337. <https://doi.org/10.1371/journal.pone.0000337>.
142. McDermott J, Bumgarner R, Samudrala R. 2005. Functional annotation from predicted protein interaction networks. *Bioinformatics* 21:3217–3226. <https://doi.org/10.1093/bioinformatics/bti514>.
143. Chua HN, Sung WK, Wong L. 2006. Exploiting indirect neighbours and topological weight to predict protein function from protein–protein interactions. *Bioinformatics* 22:1623–1630. <https://doi.org/10.1093/bioinformatics/btl145>.
144. Albert R. 2005. Scale-free networks in cell biology. *J Cell Sci* 118:4947–4957. <https://doi.org/10.1242/jcs.02714>.
145. Nguyen Ba AN, Yeh BJ, van Dyk D, Davidson AR, Andrews BJ, Weiss EL, Moses AM. 2012. Proteome-wide discovery of evolutionary conserved sequences in disordered regions. *Sci Signal* 5:rs1. <https://doi.org/10.1126/scisignal.2002515>.
146. Jeong H, Mason SP, Barabási AL, Oltvai ZN. 2001. Lethality and centrality in protein networks. *Nature* 411:41–42. <https://doi.org/10.1038/35075138>.
147. Dyer MD, Neff C, Dufford M, Rivera CG, Shattuck D, Bassaganya-Riera J, Murali TM, Sobral BW. 2010. The human-bacterial pathogen protein interaction networks of *Bacillus anthracis*, *Francisella tularensis*, and *Yersinia pestis*. *PLoS One* 5:e12089. <https://doi.org/10.1371/journal.pone.0012089>.
148. Zhou H, Gao S, Nguyen NN, Fan M, Jin J, Liu B, Zhao L, Xiong G, Tan M, Li S, Wong L. 2014. Stringent homology-based prediction of *H. sapiens*–*M. tuberculosis* H37Rv protein–protein interactions. *Biol Direct* 9:5. <https://doi.org/10.1186/1745-6150-9-5>.
149. Barabási AL, Albert R. 1999. Emergence of scaling in random networks. *Science* 286:509–512. <https://doi.org/10.1126/science.286.5439.509>.
150. Zhu X, Gerstein M, Snyder M. 2007. Getting connected: analysis and principles of biological networks. *Genes Dev* 21:1010–1024. <https://doi.org/10.1101/gad.1528707>.
151. Yook SH, Oltvai ZN, Barabási AL. 2004. Functional and topological characterization of protein interaction networks. *Proteomics* 4:928–942. <https://doi.org/10.1002/pmic.200300636>.
152. Jeong H, Tombor B, Albert R, Oltvai ZN, Barabási AL. 2000. The large-scale organization of metabolic networks. *Nature* 407:651–654. <https://doi.org/10.1038/35036627>.
153. Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain PO, Han JDJ, Chesneau A, Hao T, Goldberg DS, Li N, Martinez M, Rual JF, Lamesch P, Xu L, Tewari M, Wong SL, Zhang LV, Berriz GF, Jacotot L, Vaglio P, Reboul J, Hirozane-Kishikawa T, Li Q, Gabel HW, Elewa A, Baumgartner B, Rose DJ, Yu H, Bosak S, Sequerra R, Fraser A, Mango SE, Saxton WM, Strome S, Van Den Heuvel S, Piano F, Vandenhaute J, Sardet C, Gerstein M, Doucette-Stamm L, Gunsalus KC, Harper JW, Cusick ME, Roth FP, Hill DE, Vidal M. 2004. A map of the interactome network of the metazoan *C. elegans*. *Science* 303:540–543. <https://doi.org/10.1126/science.1091403>.
154. Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, Vijayadamar G, Pochart P, Machineni H, Welsh M, Kong Y, Zerhusen B, Malcolm R, Varrone Z, Collis A, Minto M, Burgess S, McDaniel L, Stimpson E, Spriggs F, Williams J, Neurath K, Ioime N, Agee

- M, Voss E, Furtak K, Renzulli R, Aanensen N, Carrola S, Bickelhaupt E, Lazovatsky Y, DaSilva A, Zhong J, Stanyon CA, Finley RL, White KP, Braverman M, Jarvie T, Gold S, Leach M, Knight J, Shimkets RA, McKenna MP, Chant J, Rothberg JM. 2003. A protein interaction map of *Drosophila melanogaster*. *Science* 302:1727–1736. <https://doi.org/10.1126/science.1090289>.
155. Tanaka R, Yi TM, Doyle J. 2005. Some protein interaction data do not exhibit power law statistics. *FEBS Lett* 579:5140–5144. <https://doi.org/10.1016/j.febslet.2005.08.024>.
156. Watts DJ, Strogatz SH. 1998. Collective dynamics of “small-world” networks. *Nature* 393:440–442. <https://doi.org/10.1038/30918>.
157. Amaral LA, Scala A, Barthelemy M, Stanley HE. 2000. Classes of small-world networks. *Proc Natl Acad Sci U S A* 97:11149–11152. <https://doi.org/10.1073/pnas.200327197>.
158. Fell DA, Wagner A. 2000. The small world of metabolism. *Nat Biotechnol* 18:1121–1122. <https://doi.org/10.1038/81025>.
159. Joy MP, Brock A, Ingber DE, Huang S. 2005. High-betweenness proteins in the yeast protein interaction network. *J Biomed Biotechnol* 2005:96–103. <https://doi.org/10.1155/JBB.2005.96>.
160. Newman MEJ, Girvan M. 2004. Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys* 69:e026113. <https://doi.org/10.1103/PhysRevE.69.026113>.
161. Ravasz E. 2009. Detecting hierarchical modularity in biological networks. *Methods Mol Biol* 541:145–160. https://doi.org/10.1007/978-1-59745-243-4_7.
162. Hartwell LH, Hopfield JJ, Leibler S, Murray AW. 1999. From molecular to modular cell biology. *Nature* 402:C47–C52. <https://doi.org/10.1038/35011540>.
163. Ihmels J, Friedlander G, Bergmann S, Sarig O, Ziv Y, Barkai N. 2002. Revealing modular organization in the yeast transcriptional network. *Nat Genet* 31:370–377. <https://doi.org/10.1038/ng941>.
164. Girvan M, Newman MEJ. 2002. Community structure in social and biological networks. *Proc Natl Acad Sci U S A* 99:7821–7826. <https://doi.org/10.1073/pnas.122653799>.
165. Spirin V, Mirny LA. 2003. Protein complexes and functional modules in molecular networks. *Proc Natl Acad Sci U S A* 100:12123–12128. <https://doi.org/10.1073/pnas.2032324100>.
166. Jafari M, Mirzaie M, Sadeghi M. 2015. Interlog protein network: an evolutionary benchmark of protein interaction networks for the evaluation of clustering algorithms. *BMC Bioinform* 16:319. <https://doi.org/10.1186/s12859-015-0755-1>.
167. Dunn R, Dudbridge F, Sanderson CM. 2005. The use of edge-betweenness clustering to investigate biological function in protein interaction networks. *BMC Bioinform* 6:39. <https://doi.org/10.1186/1471-2105-6-39>.
168. Chin CH, Chen SH, Ho CW, Ko MT, Lin CY. 2010. A hub-attachment based method to detect functional modules from confidence-scored protein interactions and expression profiles. *BMC Bioinform* 11:S25. <https://doi.org/10.1186/1471-2105-11-S1-S25>.
169. Wang Y, Qian X. 2014. Functional module identification in protein interaction networks by interaction patterns. *Bioinformatics* 30:81–93. <https://doi.org/10.1093/bioinformatics/btt569>.
170. Ozawa Y, Saito R, Fujimori S, Kashima H, Ishizaka M, Yanagawa H, Miyamoto-Sato E, Tomita M. 2010. Protein complex prediction via verifying and reconstructing the topology of domain–domain interactions. *BMC Bioinform* 11:350. <https://doi.org/10.1186/1471-2105-11-350>.
171. Krishnadev O, Srinivasan N. 2011. Prediction of protein–protein interactions between human host and a pathogen and its application to three pathogenic bacteria. *Int J Biol Macromol* 48:613–619. <https://doi.org/10.1016/j.ijbiomac.2011.01.030>.
172. Oliver S. 2000. Guilt-by-association goes global. *Nature* 403:601–603. <https://doi.org/10.1038/35001165>.
173. Linghu B, Snitkin ES, Holloway DT, Gustafson AM, Xia Y, DeLisi C. 2008. High-precision high-coverage functional inference from integrated data sources. *BMC Bioinform* 9:119. <https://doi.org/10.1186/1471-2105-9-119>.
174. Chua HN, Sung WK, Wong L. 2007. Using indirect protein interactions for the prediction of gene ontology functions. *BMC Bioinform* 8:S8. <https://doi.org/10.1186/1471-2105-8-S4-S8>.
175. Nabieva E, Jim K, Agarwal A, Chazelle B, Singh M. 2005. Whole-proteome prediction of protein function via graph-theoretic analysis of interaction maps. *Bioinformatics* 21:i302–i310. <https://doi.org/10.1093/bioinformatics/bti1054>.
176. Myers CL, Troyanskaya OG. 2007. Context-sensitive data integration and prediction of biological networks. *Bioinformatics* 23:2322–2330. <https://doi.org/10.1093/bioinformatics/btm332>.
177. Guo Z, Wang L, Li Y, Gong X, Yao C, Ma W, Wang D, Li Y, Zhu J, Zhang M, Yang D, Rao S, Wang J. 2007. Edge-based scoring and searching method for identifying condition-responsive protein–protein interaction sub-network. *Bioinformatics* 23:2121–2128. <https://doi.org/10.1093/bioinformatics/btm294>.
178. Zitnik M, Sosić R, Feldman MW, Leskovec J. 2019. Evolution of resilience in protein interactomes across the tree of life. *Proc Natl Acad Sci U S A* 116:4426–4433. <https://doi.org/10.1073/pnas.1818013116>.
179. Kiemer L, Cesareni G. 2007. Comparative interactomics: comparing apples and pears? *Trends Biotechnol* 25:448–454. <https://doi.org/10.1016/j.tibtech.2007.08.002>.
180. Wuchty S, Uetz P. 2014. Protein–protein interaction networks of *E. coli* and *S. cerevisiae* are similar. *Sci Rep* 4:7187. <https://doi.org/10.1038/srep07187>.
181. Berg J, Lässig M. 2004. Local graph alignment and motif search in biological networks. *Proc Natl Acad Sci U S A* 101:14689–14694. <https://doi.org/10.1073/pnas.0305199101>.
182. Kuchaiev O, Przulj N. 2011. Integrative network alignment reveals large regions of global network similarity in yeast and human. *Bioinformatics* 27:1390–1396. <https://doi.org/10.1093/bioinformatics/btr127>.
183. Aladag AE, Erten C. 2013. SPINAL: scalable protein interaction network alignment. *Bioinformatics* 29:917–924. <https://doi.org/10.1093/bioinformatics/btt071>.
184. Parida L. 2007. Discovering topological motifs using a compact notation. *J Comput Biol* 14:300–323. <https://doi.org/10.1089/cmb.2006.0142>.
185. Shao M, Yang Y, Guan J, Zhou S. 2014. Choosing appropriate models for protein–protein interaction networks: a comparison study. *Brief Bioinform* 15:823–838. <https://doi.org/10.1093/bib/bbt014>.
186. Lancichinetti A, Kivela M, Saramaki J, Fortunato S. 2010. Characterizing the community structure of complex networks. *PLoS One* 5:e11976. <https://doi.org/10.1371/journal.pone.0011976>.
187. Przulj N. 2007. Biological network comparison using graphlet degree distribution. *Bioinformatics* 23:e177–e183. <https://doi.org/10.1093/bioinformatics/btl301>.
188. Wu X, Liu Q, Jiang R. 2009. Align human interactome with phenome to identify causative genes and networks underlying disease families. *Bioinformatics* 25:98–104. <https://doi.org/10.1093/bioinformatics/btn593>.
189. Futschik ME, Chaurasia G, Herzel H. 2007. Comparison of human protein–protein interaction maps. *Bioinformatics* 23:605–611. <https://doi.org/10.1093/bioinformatics/btl683>.
190. Singh R, Xu J, Berger B. 2008. Global alignment of multiple protein interaction networks with application to functional orthology detection. *Proc Natl Acad Sci U S A* 105:12763–12768. <https://doi.org/10.1073/pnas.0806627105>.
191. Liao CS, Lu K, Baym M, Singh R, Berger B. 2009. IsoRankN: spectral methods for global alignment of multiple protein networks. *Bioinformatics* 25:i253–i258. <https://doi.org/10.1093/bioinformatics/btp203>.
192. Cai H, Lilburn TG, Hong C, Gu J, Kuang R, Wang Y. 2015. Predicting and exploring network components involved in pathogenesis in the malaria parasite via novel subnetwork alignments. *BMC Syst Biol* 9:S1. <https://doi.org/10.1186/1752-0509-9-S4-S1>.
193. Cai H, Hong C, Gu J, Lilburn TG, Kuang R, Wang Y. 2012. Module-based subnetwork alignments reveal novel transcriptional regulators in malaria parasite *Plasmodium falciparum*. *BMC Syst Biol* 6:S5. <https://doi.org/10.1186/1752-0509-6-S3-S5>.
194. Dutkowsky J, Tiuryn J. 2007. Identification of functional modules from conserved ancestral protein–protein interactions. *Bioinformatics* 23:i149–i158. <https://doi.org/10.1093/bioinformatics/btm194>.
195. Schleik S, Sun J, Raghavan B, Srnec M, Müller N, Koepfinger M, Murthy L, Zhao Z, Klein-Seetharaman J. 2012. The current *Salmonella*–host interactome. *Proteomics Clin Appl* 6:117–133. <https://doi.org/10.1002/prca.201100083>.
196. Häuser R, Blasche S, Dokland T, Haggård-Ljungquist E, von Brunn A, Salas M, Casjens S, Molineux I, Uetz P. 2012. Bacteriophage protein–protein interactions. *Adv Virus Res* 83:219–298. <https://doi.org/10.1016/B978-0-12-394438-2.00006-2>.
197. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. 2003. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* 31:258–261. <https://doi.org/10.1093/nar/gkg034>.
198. Raman K, Chandra N. 2008. *Mycobacterium tuberculosis* interactome analysis unravels potential pathways to drug resistance. *BMC Microbiol* 8:234. <https://doi.org/10.1186/1471-2180-8-234>.
199. Raman MP, Singh S, Devi PR, Velmurugan D. 2012. Uncovering potential drug targets for Tuberculosis using protein networks. *Bioinformation* 8:403–406. <https://doi.org/10.6026/97320630008403>.

200. Folador EL, de Carvalho PVSD, Silva WM, Ferreira RS, Silva A, Gromiha M, Ghosh P, Barh D, Azevedo V, Röttger R. 2016. *In silico* identification of essential proteins in *Corynebacterium pseudotuberculosis* based on protein–protein interaction networks. BMC Syst Biol 10:103. <https://doi.org/10.1186/s12918-016-0346-4>.
201. Prathiviraj R, Berchmans S, Chellapandi P. 2019. Analysis of modularity in proteome-wide protein interaction networks of *Methanothermobacter thermautotrophicus* strain ΔH and metal-loving bacteria. J Proteins Proteom 10:179–190. <https://doi.org/10.1007/s42485-019-00019-5>.
202. Bharathi M, Chellapandi P. 2019. Comparative analysis of differential proteome-wide protein–protein interaction network of *Methanobrevibacter ruminantium* M1. Biochem Biophys Rep 20:100698. <https://doi.org/10.1016/j.bbrep.2019.100698>.
203. Kumar S, Lata KS, Sharma P, Bhairappanavar SB, Soni S, Das J. 2019. Inferring pathogen–host interactions between *Leptospira interrogans* and *Homo sapiens* using network theory. Sci Rep 9:1434. <https://doi.org/10.1038/s41598-018-38329-1>.
204. Kim WY, Kang S, Kim BC, Oh J, Cho S, Bhak J, Choi JS. 2008. SynechoNET: integrated protein–protein interaction database of a model cyanobacterium *Synechocystis* sp. PCC 6803. BMC Bioinform 9:S20. <https://doi.org/10.1186/1471-2105-9-S1-S20>.
205. Cong Q, Anishchenko I, Ovchinnikov S, Baker D. 2019. Protein interaction networks revealed by proteome coevolution. Science 365:185–189. <https://doi.org/10.1126/science.aaw6718>.
206. Kong P, Huang G, Liu W. 2020. Identification of protein complexes and functional modules in *E. coli* PPI networks. BMC Microbiol 20:243. <https://doi.org/10.1186/s12866-020-01904-6>.
207. Wojcik J, Boneca IG, Legrain P. 2002. Prediction, assessment and validation of protein interaction maps in bacteria. J Mol Biol 323:763–770. [https://doi.org/10.1016/s0022-2836\(02\)01009-4](https://doi.org/10.1016/s0022-2836(02)01009-4).
208. Wojcik J, Schächter V. 2001. Protein–protein interaction map inference using interacting domain profile pairs. Bioinformatics 17:S296–S305. https://doi.org/10.1093/bioinformatics/17.suppl_1.S296.
209. Wuchty S, Müller SA, Caufield JH, Häuser R, Aloy P, Kalkhof S, Uetz P. 2018. Proteome data improves protein function prediction in the interactome of *Helicobacter pylori*. Mol Cell Proteomics 17:961–973. <https://doi.org/10.1074/mcp.RA117.000474>.
210. Park SJ, Choi JS, Kim BC, Jho SW, Ryu JW, Park D, Lee KA, Bhak J, Kim SI. 2009. PutidaNET: interactome database service and network analysis of *Pseudomonas putida* KT2440. BMC Genom 10:S18. <https://doi.org/10.1186/1471-2164-10-S3-S18>.
211. Singh A, Mai D, Kumar A, Steyn AJC. 2006. Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein–protein association. Proc Natl Acad Sci U S A 103:11346–11351. <https://doi.org/10.1073/pnas.0602817103>.
212. Zhou H, Wong L. 2011. Comparative analysis and assessment of *M. tuberculosis* H37Rv protein–protein interaction datasets. BMC Genom 12: S20. <https://doi.org/10.1186/1471-2164-12-S3-S20>.
213. Cui T, Zhang L, Wang X, He ZG. 2009. Uncovering new signaling proteins and potential drug targets through the interactome analysis of *Mycobacterium tuberculosis*. BMC Genom 10:118. <https://doi.org/10.1186/1471-2164-10-118>.
214. Edwards RA, McNair K, Faust K, Raes J, Dutilh BE. 2016. Computational approaches to predict bacteriophage–host relationships. FEMS Microbiol Rev 40:258–272. <https://doi.org/10.1093/femsre/fuv048>.
215. Dutilh BE, Cassman N, McNair K, Sanchez SE, Silva GGZ, Boling L, Barr JJ, Speth DR, Seguritan V, Aziz RK, Felts B, Dinsdale EA, Mokili JL, Edwards RA. 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat Commun 5:4498. <https://doi.org/10.1038/ncomms5498>.
216. Shapiro JW, Putonti C. 2018. Gene co-occurrence networks reflect bacteriophage ecology and evolution. mBio 9:e01870–17. <https://doi.org/10.1128/mBio.01870-17>.
217. Lima-Mendez G, Van Helden J, Toussaint A, Leplae R. 2008. Reticulate representation of evolutionary and functional relationships between phage genomes. Mol Biol Evol 25:762–777. <https://doi.org/10.1093/molbev/msn023>.
218. Paez-Espino D, Eloë-Fadrosch EA, Pavlopoulos GA, Thomas AD, Huntemann M, Mikhailova N, Rubin E, Ivanova NN, Kyrpides NC. 2016. Uncovering Earth's virome. Nature 536:425–430. <https://doi.org/10.1038/nature19094>.
219. Wang W, Ren J, Tang K, Dart E, Ignacio-Espinoza JC, Fuhrman JA, Braun J, Sun F, Ahlgren NA. 2020. A network-based integrated framework for predicting virus–prokaryote interactions. NAR Genom Bioinform 2: lqaa044. <https://doi.org/10.1093/nargab/lqaa044>.
220. Métris A, Sudhakar P, Fazekas D, Demeter A, Ari E, Olbei M, Branchu P, Kingsley RA, Baranyi J, Korcsmáros T. 2017. SalmoNet, an integrated network of ten *Salmonella enterica* strains reveals common and distinct pathways to host adaptation. NPJ Syst Biol Appl 3:31. <https://doi.org/10.1038/s41540-017-0034-z>.
221. Loaiza CD, Duhan N, Lister M, Kaundal R. 2020. *In silico* prediction of host–pathogen protein interactions in melioidosis pathogen *Burkholderia pseudomallei* and human reveals novel virulence factors and their targets. Brief Bioinform 22:bbz162.
222. Durmuş Tekir SD, Ülgen KO. 2013. Systems biology of pathogen–host interaction: networks of protein–protein interaction within pathogens and pathogen–human interactions in the post-genomic era. Biotechnol J 8:85–96. <https://doi.org/10.1002/biot.201200110>.
223. Pan A, Lahiri C, Rajendiran A, Shanmugham B. 2016. Computational analysis of protein interaction networks for infectious diseases. Brief Bioinform 17:517–526. <https://doi.org/10.1093/bib/bbv059>.
224. Mahajan G, Mande SC. 2017. Using structural knowledge in the protein data bank to inform the search for potential host–microbe protein interactions in sequence space: application to *Mycobacterium tuberculosis*. BMC Bioinform 18:201. <https://doi.org/10.1186/s12859-017-1550-y>.
225. Zanzoni A, Spinelli L, Braham S, Brun C. 2017. Perturbed human sub-networks by *Fusobacterium nucleatum* candidate virulence proteins. Microbiome 5:89. <https://doi.org/10.1186/s40168-017-0307-1>.
226. Acharya D, Dutta T. 2021. Elucidating the network features and evolutionary attributes of intra- and interspecific protein–protein interactions between human and pathogenic bacteria. Sci Rep 11:190. <https://doi.org/10.1038/s41598-020-80549-x>.
227. Li ZG, He F, Zhang Z, Peng YL. 2012. Prediction of protein–protein interactions between *Ralstonia solanacearum* and *Arabidopsis thaliana*. Amino Acids 42:2363–2371. <https://doi.org/10.1007/s00726-011-0978-z>.
228. Sahu SS, Weirick T, Kaundal R. 2014. Predicting genome-scale *Arabidopsis–Pseudomonas syringae* interactome using domain and interolog-based approaches. BMC Bioinform 15:S13. <https://doi.org/10.1186/1471-2105-15-S11-S13>.
229. Zhang L, Liu JY, Gu H, Du Y, Zuo JF, Zhang Z, Zhang M, Li P, Dunwell JM, Cao Y, Zhang Z, Zhang YM. 2018. *Bradyrhizobium diazoefficiens* USDA 110- Glycine max interactome provides candidate proteins associated with symbiosis. J Proteome Res 17:3061–3074. <https://doi.org/10.1021/acs.jproteome.8b00209>.
230. Evans P, Dampier W, Ungar L, Tozeren A. 2009. Prediction of HIV-1 virus–host protein interactions using virus and host sequence motifs. BMC Med Genom 2:27. <https://doi.org/10.1186/1755-8794-2-27>.
231. Becerra A, Bucheli VA, Moreno PA. 2017. Prediction of virus–host protein–protein interactions mediated by short linear motifs. BMC Bioinform 18:163. <https://doi.org/10.1186/s12859-017-1570-7>.
232. Sarmady M, Dampier W, Tozeren A. 2011. Sequence- and interactome-based prediction of viral protein hotspots targeting host proteins: a case study for HIV Nef. PLoS One 6:e20735. <https://doi.org/10.1371/journal.pone.0020735>.
233. Emamjomeh A, Goliaei B, Zahiri J, Ebrahimpour R. 2014. Predicting protein–protein interactions between human and hepatitis C virus via an ensemble learning method. Mol Biosyst 10:3147–3154. <https://doi.org/10.1039/c4mb00410h>.
234. Cui G, Fang C, Han K. 2012. Prediction of protein–protein interactions between viruses and human by an SVM model. BMC Bioinform 13:S5. <https://doi.org/10.1186/1471-2105-13-S7-S5>.
235. Halder AK, Dutta P, Kundu M, Basu S, Nasipuri M. 2018. Review of computational methods for virus–host protein interaction prediction: a case study on novel Ebola–human interactions. Brief Funct Genom 17:381–391. <https://doi.org/10.1093/bfgp/elix026>.
236. Li Y, Zhang Z, Yang L, Lian X, Xie Y, Li S, Xin S, Cao P, Lu J. 2020. The MERS-CoV receptor DPP4 as a candidate binding target of the SARS-CoV-2 spike. iScience 23:101160. <https://doi.org/10.1016/j.isci.2020.101160>.
237. Dey L, Chakraborty S, Mukhopadhyay A. 2020. Machine learning techniques for sequence-based prediction of viral–host interactions between SARS-CoV-2 and human proteins. Biomed J 43:438–450. <https://doi.org/10.1016/j.bj.2020.08.003>.
238. Lu H, Zhou Q, He J, Jiang Z, Peng C, Tong R, Shi J. 2020. Recent advances in the development of protein–protein interactions modulators: mechanisms and clinical trials. Signal Transduct Target Ther 5:213. <https://doi.org/10.1038/s41392-020-00315-3>.
239. Rapanoel HA, Mazandu GK, Mulder NJ. 2013. Predicting and analyzing interactions between *Mycobacterium tuberculosis* and its human host. PLoS One 8:e67472. <https://doi.org/10.1371/journal.pone.0067472>.

240. Mei S. 2018. *In silico* enhancing *M. tuberculosis* protein interaction networks in STRING to predict drug-resistance pathways and pharmacological risks. *J Proteome Res* 17:1749–1760. <https://doi.org/10.1021/acs.jproteome.7b00702>.
241. Uddin R, Tariq SS, Azam SS, Wadood A, Moin ST. 2017. Identification of histone deacetylase (HDAC) as a drug target against MRSA via interolog method of protein–protein interaction prediction. *Eur J Pharm Sci* 106:198–211. <https://doi.org/10.1016/j.ejps.2017.06.003>.
242. Li S, Su Z, Zhang C, Xu Z, Chang X, Zhu J, Xiao R, Li L, Zhou R. 2018. Identification of drug target candidates of the swine pathogen *Actinobacillus pleuropneumoniae* by construction of protein–protein interaction network. *Genes Genomics* 40:847–856. <https://doi.org/10.1007/s13258-018-0691-3>.
243. Kingkaw A, Nakphaichit M, Suratannon N, Nitisinprasert S, Wongoutong C, Chatchatee P, Krobthong S, Charoenlappanit S, Roytrakul S, Vongsangnak W. 2020. Analysis of the infant gut microbiome reveals metabolic functional roles associated with healthy infants and infants with atopic dermatitis using metaproteomics. *PeerJ* 8:e9988. <https://doi.org/10.7717/peerj.9988>.
244. Typas A, Sourjik V. 2015. Bacterial protein networks: properties and functions. *Nat Rev Microbiol* 13:559–572. <https://doi.org/10.1038/nrmicro3508>.
245. Skinnider MA, Scott NE, Prudova A, Stoykov N, Stacey RG, Gsponer J, Foster L. 2016. An atlas of protein–protein interactions across mammalian tissues. *SSRN*. <https://doi.org/10.2139/ssrn.3219264>.
246. Folador EL, Hassan SS, Lemke N, Barh D, Silva A, Ferreira RS, Azevedo V. 2014. An improved interolog mapping-based computational prediction of protein–protein interactions with increased network coverage. *Integr Biol (Camb)* 6:1080–1087. <https://doi.org/10.1039/c4ib00136b>.
247. Brown KR, Jurisica I. 2007. Unequal evolutionary conservation of human protein interactions in interologous networks. *Genome Biol* 8:R95. <https://doi.org/10.1186/gb-2007-8-5-r95>.
248. Das J, Vo TV, Wei X, Mellor JC, Tong V, Degatano AG, Wang X, Wang L, Cordero NA, Kruer-Zerhusen N, Matsuyama A, Pleiss JA, Lipkin SM, Yoshida M, Roth FP, Yu H. 2013. Cross-species protein interactome mapping reveals species-specific wiring of stress response pathways. *Sci Signal* 6:ra38. <https://doi.org/10.1126/scisignal.2003350>.
249. Vo TV, Das J, Meyer MJ, Cordero NA, Akturk N, Wei X, Fair BJ, Degatano AG, Fragoza R, Liu LG, Matsuyama A, Trickey M, Horibata S, Grimson A, Yamano H, Yoshida M, Roth FP, Pleiss JA, Xia Y, Yu H. 2016. A proteome-wide fission yeast interactome reveals network evolution principles from yeasts to human. *Cell* 164:310–323. <https://doi.org/10.1016/j.cell.2015.11.037>.
250. Sun J, Li Y, Zhao Z. 2007. Phylogenetic profiles for the prediction of protein–protein interactions: how to select reference organisms? *Biochem Biophys Res Commun* 353:985–991. <https://doi.org/10.1016/j.bbrc.2006.12.146>.
251. Lemmens I, Lievens S, Tavernier J. 2010. Strategies towards high-quality binary protein interactome maps. *J Proteomics* 73:1415–1420. <https://doi.org/10.1016/j.jprot.2010.02.001>.
252. Yu J, Finley RL. 2009. Combining multiple positive training sets to generate confidence scores for protein–protein interactions. *Bioinformatics* 25:105–111. <https://doi.org/10.1093/bioinformatics/btn597>.
253. Braun P. 2012. Interactome mapping for analysis of complex phenotypes: insights from benchmarking binary interaction assays. *Proteomics* 12:1499–1518. <https://doi.org/10.1002/pmic.201100598>.
254. Shatsky M, Allen S, Gold BL, Liu NL, Juba TR, Revoco SA, Elias DA, Prathapam R, He J, Yang W, Szakal ED, Liu H, Singer ME, Geller JT, Lam BR, Saini A, Trotter VV, Hall SC, Fisher SJ, Brenner SE, Chhabra SR, Hazen TC, Wall JD, Witkowska HE, Biggin MD, Chandonia JM, Butland G. 2016. Bacterial interactomes: interacting protein partners share similar function and are validated in independent assays more frequently than previously reported. *Mol Cell Proteomics* 15:1539–1555. <https://doi.org/10.1074/mcp.M115.054692>.
255. Bouveret E, Brun C. 2012. Bacterial interactomes: from interactions to networks. *Methods Mol Biol* 804:15–33. https://doi.org/10.1007/978-1-61779-361-5_2.
256. Nicod C, Banaei-Esfahani A, Collins BC. 2017. Elucidation of host–pathogen protein–protein interactions to uncover mechanisms of host cell rewiring. *Curr Opin Microbiol* 39:7–15. <https://doi.org/10.1016/j.mib.2017.07.005>.
257. Caufield JH, Wimble C, Shary S, Wuchty S, Uetz P. 2017. Bacterial protein meta-interactomes predict cross-species interactions and protein function. *BMC Bioinform* 18:171. <https://doi.org/10.1186/s12859-017-1585-0>.
258. Carro L. 2018. Protein–protein interactions in bacteria: a promising and challenging avenue towards the discovery of new antibiotics. *Beilstein J Org Chem* 14:2881–2896. <https://doi.org/10.3762/bjoc.14.267>.
259. Olson EJ, Tabor JJ. 2012. Post-translational tools expand the scope of synthetic biology. *Curr Opin Chem Biol* 16:300–306. <https://doi.org/10.1016/j.cbpa.2012.06.003>.
260. Guo J, Li H, Chang JW, Lei Y, Li S, Chen LL. 2013. Prediction and characterization of protein–protein interaction network in *Xanthomonas oryzae* pv. *oryzae* PXO99 A. *Res Microbiol* 164:1035–1044. <https://doi.org/10.1016/j.resmic.2013.09.001>.