 doi:10.20944/preprints202110.0185.v1

Expanding interactome analyses beyond model eukaryotes

Katherine James*[1], Anil Wipat[2], and Simon J. Cockell[3]

[1]Applied Sciences, Northumbria University, Newcastle upon Tyne, UK.
[2]School of Computing, Newcastle University, Newcastle upon Tyne, UK.
[3]School of Biomedical, Nutritional and Sports Sciences, Newcastle University, Newcastle upon Tyne, UK.

Abstract

Interactome analyses have traditionally been applied to yeast, human and other model organisms due to the availability of protein-protein interactions data for these species. Recently these techniques have been applied to more diverse species using computational interaction prediction from genome sequence and other data types. This review describes the various types of computational interactome networks that can be created and how they have been used in diverse eukaryotic species, highlighting some of the key interactome studies in non-model organisms.

Key Words: Systems biology; Integrative bioinformatics; Interactomics; Eukaryotes; Non-models; Network biology.

Introduction

In the early 2000s technologies were developed for high-throughput (HTP) production of genome-wide genetic [1–3] and physical interaction data [4–7], allowing genes and their products to be studied in the context of whole cellular systems [8]. Systems approaches can be used to understand the link between genotypes and phenotypes [9, 10], the evolution of gene function [11–13], and how cellular biology changes during disease [14, 15].

One of the fundamental goals of these analyses is verification of the cellular interactome [16, 17]. The interactome is vital to understanding cellular biology [18] since many biological functions can only be understood as part of the spatio-temporal interactions of the cell [19, 20]. Defining the interactome is not straightforward since the cell contains other molecules that interact with proteins [19, 21, 22]. The associations between genes/proteins can be functional rather than a direct physical interaction, for instance shared complex membership [23–25]. Here the interactome is defined loosely as "*the entire complement of functional molecular associations that may occur in a cell*" in order to encompass the range of networks discussed.

Interactomes are often represented as networks (graphs) [26], allowing both visual and computational analysis of their structure and connectivity [27–31]. Graph theoretic analyses allows interactome data to be used in a number of ways: detection of protein complexes [32–34]; prediction of protein functions [35–37]; identification of evolutionary relationships [38–41]; and inference

*Corresponding author katherine.p.m.james@northumbria.ac.uk

 doi:10.20944/preprints202110.0185.v1

of novel interactions that were not detected experimentally [42, 43]. These analyses have been extensively used in yeast and human due to the availability of experimental data for these species. Recently experimental techniques have been applied to more diverse model species and interactomes have been computationally-predicted in non-model organisms using experimental data from other species. Here, we describe the types of interactome networks that can be created and how they have been used in diverse species, highlighting some of the key studies in non-model organisms.

Predicted interactomes

The detection of the physical binding of proteins, either binary interaction or complex co-membership [44], is the basis for the majority of interactome analyses. Several PPI detection methods, including yeast two-hybrid (Y2H) and affinity purification combined with mass spectometry (AP-MS), have been used to create large-scale interactomes in yeast, humans and other model organiams [4–7, 45–51]. Several other methods, such as co-immunoprecipitation (Co-IP) [52], co-fractionation [53], and cross-linking [54], can also be used to detect interaction. Co-IP is often used to confirm interactions detected by Y2H and AP-MS.

In addition, genetic interactions (GIs) can be detected using a number of methods including RNA interference (RNAi) screens [55], synthetic genetic arrays (SGA) [1], heterozygous diploid-based synthetic lethality analysis on microarrays (dSLAM) [2], and epistatic mini-array profiles (E-MAPs) [3]. Unlike PPI, GIs connect genes with related function, but which protein products are less likely to have a physical interaction [56]; genetically-interacting protein pairs are commonly components of the same pathway or complex and have a relatively high level of conservation across species [57].

Experimental determination of PPI networks is time-consuming, expensive and not currently possible for all species. Therefore, a number of computational methods for the prediction of PPIs have been developed over the past decade and are described in the next sections.

Interologs

Patterns of protein-protein interaction are conserved [58–60] and interactomes share a set of topological rules [11]. Network 'hubs' (highly-connected proteins) are conserved and often essential [30, 61–63] with slower evolutionary rates and conserved sequences [39, 62, 64–66]. Hubs and key cellular interactions are detectable between eukaryotic species [59], and also between eukaryotes and prokaryotes [40, 58, 67].

Conserved interactions, termed 'interologs' (direct), 'associologs' (functional) or 'regulogs' (regulatory), can therefore be transferred between species [68–74] providing systems-level network analysis in organisms that lack empirical interaction data (Figure 1). Co-expression patterns are also conserved between species [75, 76] allowing microarray and RNA-seq data to be used to predict functional links between species in the same way [75, 77].

Domain-domain and structural interactions

The conservation of interacting domains can be used to infer interactions since these domains are crucial to proteins' roles and more highly conserved [78–80]; the presence of specific domains in pairs of proteins can be indicative of PPI [79, 81–84]. Protein physiochemical properties, such as charge, hydrophobicity and structure, can be used in combination with domain and sequence to infer interaction [85–87]. The domain-domain interaction (DDI) method is often used in conjunction with interolog mapping [61, 88–92].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

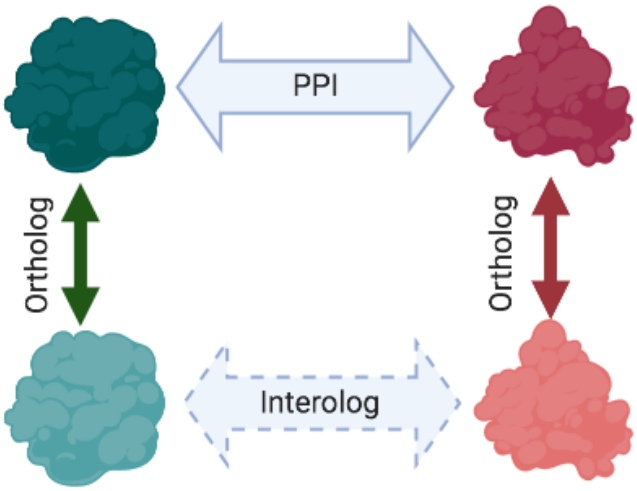

Figure 1: Interologs are mapped from known PPIs via orthology.

The 3D structure of a protein forms an active site essential for its function. It is possible to predict interactions based on these structures, for instance by docking or threading methods [93–103], and therefore to predict interactions in species lacking experimental interactome data [104]. Methods that combine sequence and structural data can improve prediction compared with sequence or structure alone [86, 105, 106]. The *in silico* two-hybrid method (i2H) takes advantage of sequence and structural conservation to predict PPI by comparison of multiple sequence alignments to identify correlated mutations [107].

Genome sequence-based interactions

Several aspects of genomic sequence can be used to predict PPI [108]. Gene context (GC) can indicate interaction since interacting proteins have increased conservation of their gene order in comparison to non-interacting proteins [109–111]; this prediction method is most accurate in bacterial genomes due to their operons [112], although conservation of gene order has also been observed in mammals [113]. Gene fusion (GF) events can reveal protein interaction since proteins that are fused as a single entity in one species are likely to have a functional link in other species in which they are encoded separately [114–116]. The distribution of gene sequences across species, termed their phylogenetic profile, is also conserved for many interacting pairs [38, 117–122] since they evolve at the same rate [39]. Two genes that have similar profiles are likely to have co-evolved and their products may interact physically or have shared function [123–127].

Integrated interactomes

The number of interactions common to different experimental datasets can be surprisingly low [33, 128–131] because each individual experiments only measure certain aspects of the cell's behaviour and the resulting datasets are incomplete [132, 133]. A more complete view of the interactome can be produced by incorporating multiple sources of interaction evidence [134] (Figure 2 A). This approach reduces the impact of experimental noise in HTP datasets [23, 135–137] and reveals global properties not evident in a single data type [19]. Integrated networks have advanced our understanding of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

several areas including cellular biology [23, 138], disease processes [14] and evolution [11, 13, 139].

Early interactome studies only linked proteins which had physical interactions (either binary or complex) [4, 5]. In functional integrated networks, pairs of proteins/genes are linked if they have any type of association; links may represent the gene/proteins' involvement in the same cellular processes without direct interaction [20], for instance via genetic interactions [56], co-localisation [140, 141], or co-citation [142–145]. The greater density of links provided by functional data provides a more informative basis for network analysis than physical interactions alone. Functional networks have been used to analyse data from yeast [146, 147] and human [148], and to compare patterns of interaction across multiple species [149].

Probabilistic networks and machine learning

At the simplest level datasets can be combined naïvely into a network in which nodes represent genes or gene products, and edges represent any type of functional interaction between them [150–153] (Figure 2 A). Such networks are useful for the basic visualisation of integrated results but no attention is paid to the amount of evidence for each interaction.

A more useful network can use interaction weights to represent the number of lines of evidence for each interaction (Figure 2 B). This weighting provides a measure of confidence since interactions with several sources of evidence are more likely to be true interactions [44, 112, 136, 151, 153–159]. Taking these evidence levels into account allows thresholding of networks to produce high quality sub-networks that are supported by multiple lines of evidence (Figure 2 D) [153, 160–162].

However, the quality of different datasets, in terms of coverage of the genome and accuracy, depends upon the experimental technique used. Probabilistic functional integrated networks (PFINs) take data quality into account by assessing datasets' quality prior to integration (Figure 2 C) [23, 25, 147]. The confidence scores are produced by statistical comparison with a Gold Standard dataset [163–165]: a high-confidence set of interactions believed to be biologically correct [166, 167]. This benchmarking reduces noise from HTP datasets, produces consistent integration of interactions from different studies, and allows the use of thresholding (Figure 2 E) and statistical algorithms that take these probabilities into account [168]. Probabilistic networks have been created for yeast and a number of other species using a variety of methods and Gold Standards. These networks can then be used to detect protein complexes [32, 169–171], annotate proteins [37, 159, 172] and predict missing interactions [173, 174].

Finally, machine learning approaches can predict PPI using a variety of high quality data types during classifier training [175–181]. The use of multiple data types improves prediction accuracy over single-type prediction methods [182, 183]. Popular algorithms for PPI prediction include support vector machines [85, 176, 184, 185], random forest algorithms [176, 182, 186–188], and naïve Bayes [176, 189].

Computational interactome networks of diverse species

Fungi

While the model budding yeast *Saccharomyces cerevisiae* was the first species to be used in high-throughput interaction screens [1, 2, 4–7], the model fission yeast *Schizosaccharomyces pombe* has lagged behind, with few HTP datasets being produced until relatively recently [190–192]. Comparison of the stress-response interactome of *S. pombe*, StressNet, with the network from *S. cerevisiae* indicated that most stress-related interactions are not conserved and have undergone considerable rewiring during evolution [193]. Binary interactions of *S. pombe* were found to be better conserved

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

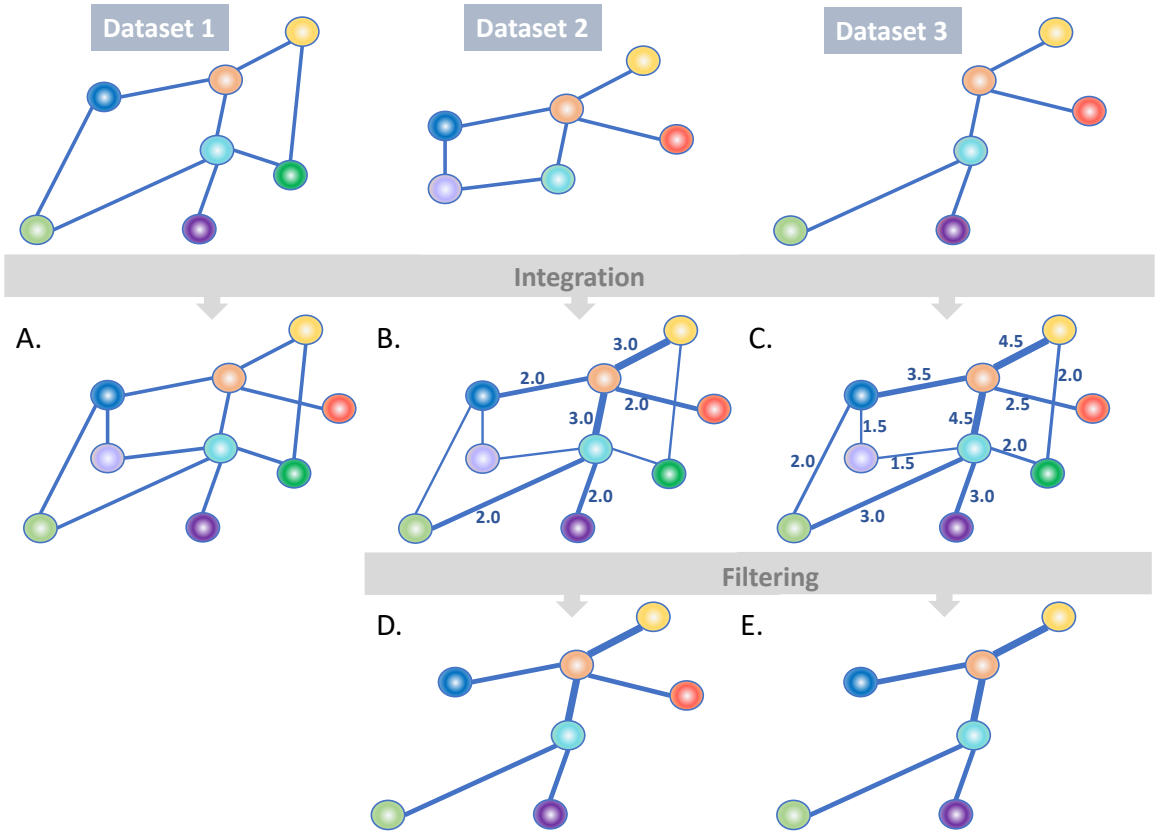

Figure 2: Integration strategies for three datasets measuring interaction (blue lines) between nine proteins (coloured circles). A. The union of the three datasets is produced. B. The interactions are annotated with the number of datasets with evidence for that interaction. C. The datasets have confidence scores (here A = 2.0, B = 1.5, C = 1.0) and the interactions are annotated with the sum of these scores. D. Network B can be filtered to select interactions with > 1 lines of evidence. E. Interaction confidence scores can be filtered to select interactions with high confidence (here $>=3.0$). In this case the interaction between the pink and red proteins is lost since it does not pass the threshold despite having two lines of evidence in D.

with humans than with budding yeast, further supporting this evolutionary rewiring [191]. Genetic interactions appear to have higher conservation between *S. pombe* and budding yeast [194], with $\sim 30\%$ of synthetic lethal interactions found to be conserved [57].

Mass spectroscopy has been used in *S. cerevisiae*, *Candida albicans*, and *S. pombe* to quantify the evolutionary rate of change in phosphorylation, indicating that kinases have a lower rate of conservation, and several of the *S. pombe* results were confirmed using E-MAPs [195]. A more recent E-MAP study suggested a hierarchical model for the evolution of genetic interactions [196]. Despite the evidence for evolutionary rewiring of *S. pombe* PPIs, it is possible to use *S. cerevisiae* and other data to computationally predict *S. pombe* interactions [188, 197, 198]. In particular the Pombe Interactome was predicted using over 100 protein features by comparison with budding yeast and several predicted interactions for the Cbf11 transcription factor were subsequently confirmed

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

using AP-MS [199].

Table 1: Computational interactomes in *Candida albicans* and other budding yeasts.

Species	Methodology	Proteins	Interactions	Ref
Candida albicans	Literature curated	-	1,208	[200]
	Machine learning	284	398	[188]
Candida dubliniensis	Integrated	1,272	4,699	[13]
Candida glabrata	Machine learning	2,005	5,089	[188]
	Integrated	1,423	5,421	[13]
Candida tropicalis	Integrated	1,305	4,941	[13]
Clavispora lusitaniae	Integrated	1,254	4,524	[13]
Debaryomyces hansenii	Machine learning	1,229	2,338	[188]
Eremothecium gossypii	Integrated	1,346	4,757	[13]
	Machine learning	1,014	1,971	[188]
Eremothecium cymbalariae	Integrated	1,087	3,441	[13]
Kluyveromyces lactis	Integrated	1,437	5,636	[13]
	Machine learning	1,658	38,20	[188]
Komagataella phaffii	Integrated	1,160	3,948	[13]
Lachancea thermotolerans	Integrated	1,264	4,531	[13]
Lodderomyces elongisporus	Integrated	1,225	4,284	[13]
	Machine learning	1,024	1,914	[188]
Meyerozyma guilliermondii	Integrated	1,313	4,964	[13]
Naumovozyma castellii	Integrated	1,432	5,197	[13]
Naumovozyma dairenensis	Integrated	1,424	5,289	[13]
Scheffersomyces stipitis	Integrated	1,323	5,230	[13]
	Machine learning	1,000	1,754	[188]
Tetrapisispora phaffii	Integrated	1,376	4,848	[13]
Torulaspora delbrueckii	Integrated	1,246	4,395	[13]
Vanderwaltozyma polyspora	Integrated	1,294	4,349	[13]
Yarrowia lipolytica	Integrated	1,264	4,605	[13]
	Machine learning	1,072	1,810	[188]
Zygosaccharomyces rouxii	Integrated	1,250	4,265	[13]

Candida albicans is a budding yeast and opportunistic pathogen found in the human gut [201]. Production of interaction data in this species has lagged behind other yeasts due to non-standard codon usage [202], which requires modified methodologies [203–205]. Schoeters and colleagues have provided a manually curated list *C. albicans* protein-protein interaction data [200] and several other interactome datasets have been produced in this and other budding yeasts (Table 1). Ascomycetes and other Fungi have also been investigated using computational network prediction (Tables 2 and 3). The PHI-Nets resource contains networks, produced using interologs and DDI mapping from *S. cerevisiae* and *S. pombe*, for fifteen Ascomycetes [206] including the pathogenic fungi *Aspergillus fumigatus* [207] and *Fusarium graminearum* [208]. The FPPI database is a resource for *F. graminearum*, which provides confidence scored interologs covering $\sim$ 52% of the proteome [88]. In a later study, FPPI interactions were combined with gene expression data to predict a subnetwork of pathogenicity genes; two interconnected network modules were identified that were enriched in G-protein coupled receptors and MAPK signalling pathways, and which contained several known pathogenicity genes [209]. Finally, a large-scale study by Zitnik and colleagues produced computa-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

tionally predicted networks for over 40 fungi [13].

Combining networks from pathogenic yeasts and their hosts aids understanding of infection processes and can identify potential drug targets. Microarray data from *C. albicans* infected zebrafish were used to predict networks at different stages of infection and decipher the mechanisms underlying *C. albicans* pathogenicity [210]. Time-course *C. albicans*-zebrafish transcriptomics mapped to interolog networks indicated that redox status is crucial to infection in this species [211]. Similarly, comparison of the early and late stages of *C. albicans* infection identified important functional modules in both the pathogenic and defensive mechanisms [212]. Remmele and co-workers used interolog mapping to identify host-pathogen interactions of *A. fumigatus* during human and mouse infection, highlighting the roles of the PLB1 virulence factor and APP anti-fungal host protein [213]. Several network studies have also investigated yeast infection in plants [61, 89, 91, 214–217].

Table 2: Computational interactomes in Ascomycetes yeasts. DDI: domain-domain interactions; PFIN: probabilistic functional integrated network.

Species	**Methodology**	**Proteins**	**Interactions**	**Ref**
Aspergillus clavatus	Integrated	1,479	6,747	[13]
Aspergillus fischeri	Integrated	1,577	8,010	[13]
Aspergillus flavus	Integrated	1,668	9,044	[13]
Aspergillus fumigatus	Interolog, DDI	5,925	277,441	[206]
	Integrated	1,537	8,028	[13]
	Machine learning	813	1,300	[188]
Aspergillus oryzae	Integrated	1,400	6,125	[13]
Aspergillus nidulans	Integrated	1,459	6,792	[13]
	Machine learning	921	1,521	[188]
Aspergillus niger	Integrated	1,279	5,391	[13]
Bipolaris sorokiniana	Interolog, DDI	5,389	264,403	[206]
Blumeria graminis	Interolog, DDI	3,816	154,218	[206]
Botrytis cinerea	Interolog, DDI	6,416	344,586	[206]
	Integrated	1,315	5,028	[13]
	Machine learning	852	1,387	[188]
Coccidioides immitis	Integrated	1,193	4,448	[13]
Coccidioides posadasii	Integrated	1,228	4,638	[13]
Colletotrichum gloeosporioides	Interolog, DDI	8,161	444,775	[206]
Colletotrichum graminicola	Interolog, DDI	6,514	297,282	[206]
Fusarium graminearum	PFIN	7,406	223,166	[88]
	PFIN, gene expression	127	259	[209]
	Interolog, DDI	7,062	381,518	[206]
Fusarium oxysporum	Interolog, DDI	8,292	452,631	[206]
Fusarium solani	Integrated	1,842	12,418	[13]
Fusarium verticillioides	Interolog, DDI	7,094	334,015	[206]
Histoplasma capsulatum	Integrated	1,200	4,358	[13]
Leptosphaeria maculans	Interolog, DDI	5,327	221,687	[206]
Magnaporthe (Pyricularia) grisea	Interolog	3,017	11,674	[61]
	Machine learning	629	953	[188]
Magnaporthe (Pyricularia) oryzae	Interolog, DDI	6,071	287,159	[206]

Continued on next page

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

Table 2 – continued from previous page

Species	Methodology	Proteins	Interactions	Ref
	Integrated	1,304	5,028	[13]
Neurospora crassa	Interolog, DDI	3,006	27,588	[218]
	Integrated	1,364	5,028	[13]
	Machine learning	1,161	2,186	[188]
Paracoccidioides lutzii	Integrated	1,124	3,963	[13]
Parastagonospora nodorum	Integrated	1,499	6,759	[13]
	Machine learning	255	250	[188]
Phomopsis longicolla	Interolog	3,868	215,255	[217]
Penicillium rubens	Integrated	1,597	7,358	[13]
Podospora anserina	Integrated	1,423	5,847	[13]
Pyrenophora teres	Integrated	1,450	6,372	[13]
Sclerotinia sclerotiorum	Interolog, DDI	3,803	118,987	[206]
	Integrated	856	2,400	[13]
Sordaria macrospora	Integrated	1,333	5,270	[13]
Thermothelomyces thermophilus	Integrated	1,339	5,216	[13]
Thermothielavioides terrestris	Integrated	1,357	5,171	[13]
Trichoderma reesei	Integrated	1,336	5,194	[13]
Trichophyton benhamiae	Integrated	1,206	4,366	[13]
Trichophyton rubrum	Integrated	186	1,139	[216]
Trichophyton verrucosum	Integrated	1,206	4,567	[13]
Tuber melanosporum	Integrated	1,135	3,907	[13]
Uncinocarpus reesii	Integrated	1,153	4,422	[13]
Ustilaginoidea virens	Interolog, DDI	3,305	20,217	[89]
	Machine learning	604	943	[188]
Verticillium alfalfae	Integrated	1,421	6,182	[13]
Verticillium dahliae	Interolog, DDI	5,801	247,581	[206]
Zymoseptoria tritici	Interolog, DDI	5,609	251,215	[206]

Plants

The plant *Arabidopsis thaliana*, thale cress, was chosen as the first 'model' plant due to its small genome since and diploid nature that made genetic manipulation relatively simple [219]. *A. thaliana* has most well-characterised interactome of the plants with several experimental [220–227] and computational datasets [13, 149, 187, 188, 198, 228–238], many of which form the basis of interactome studies in other plants [238–243]. Cross-species interactomes have also investigated infection in this species [215, 244].

More recently, new plant models have been developed [219] and several interactome studies have been carried out in other plant species, many of which have commercial value (Table 4). Unlike *A. thaliana*, very little interaction data have been produced for rice, *Oryza sativa*, despite its economic importance as a staple food. The majority of interactome networks in this species were computationally-predicted, although some experimental data have been produced [245–247]. Interolog mapping was used to create a proteome-wide interactome for rice, which correlated well

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

Table 3: Computational interactomes in other yeasts. DDI: domain-domain interactions.

Species	Methodology	Proteins	Interactions	Ref
Coprinopsis cinerea	Integrated	1,396	5,436	[13]
Cryptococcus neoformans	Integrated	1,134	3,986	[13]
	Machine learning	1,006	2,004	[188]
Cryptococcus gattii	Integrated	1,130	4,019	[13]
Encephalitozoon intestinalis	Integrated	246	591	[13]
Encephalitozoon hellem	Integrated	254	607	[13]
Encephalitozoon cuniculi	Integrated	265	574	[13]
	Machine learning	154	200	[188]
Laccaria bicolor	Integrated	1,384	5,507	[13]
Malassezia globosa	Integrated	907	2,636	[13]
Puccinia graminis	Integrated	1,353	4,449	[13]
Rhizoctonia solani	Interolog, DDI	1,773	6,705	[91]
Schizophyllum commune	Integrated	1,405	5,382	[13]

with co-expression data [248]. RicePPINet used structural and functional information as machine learning inputs to predict the rice interactome and the resulting network was used to identify genes involved in disease resistance and drought tolerance [249]. PlaPPISite comprises 36,420 interologs predicted using several computational methods [238], while the predicted Rice Interactome Network, PRIN, contains 76,585 interologs [250, 251]. RiceNet is a PFIN for rice produced using data from several model organisms [252]; an updated RiceNet network has 1,775,000 interactions between 25,765 genes [253].

BarleyNet is a PFIN for barley, *Hordeum vulgare*, based on orthology to *Arabidopsis* and rice [254]. Five computationally-predicted interactomes were produced for *Zea mays*, maize [13, 231, 238, 239, 255]. In horse gram, *Macrotyloma uniflorum*, an interolog network of over 6,000 interactions has been produced [256]. Predicted interactomes have also been created for many other plant species (Table 4). For instance, PTIR is a tomato interolog network based on orthology to six model organisms containing over 12,000 high confidence interactions, ten of which were experimentally verified [257]. A turnip (*Brassica rapa*) interolog network has been inferred from *A. thalina* data [258] and a predicted interactome was produced for coffee, *Coffea canephora*, using orthology to model organisms including *A. thalaina* [259]. Predicted networks have also been built for the tea plant *Camellia sinensis* [92, 260], cassava *Manihot esculenta* [261, 262], orange *Citrus sinensis* [263], thai basil, *Ocimum tenuiflorum* [264], the poplar tree *Populus trichocarpa*, and the moss *Physcomitrella patens* [187].

Compuatational methods have been used to compare networks between multiple plant species. For instance, PlaPPISite contains interolog networks for eleven other plant species inferred using structural and domain information which provides accurate prediction of interaction sites in these species [238]. Ding and colleagues used functional interaction data to infer confidence-scored interolog networks for three species—*A. thaliana*, *Glycine max* (soybean) and *Z. mays*—many interactions of which were supported by literature curated evidence [231]. Vandereyken and co-workers compared hub proteins between major plant interactome studies and found that many are involved in stress responses [265]. Zitnik and colleagues' large-scale interactome study created networks for sixteen plant species [13]. Finally, several studies have produced plant-pathogen networks [89, 266]. Fungal-plant networks are discussed above and reviewed in [267].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

Table 4: Computational interactomes in plants. DDI: domain-domain interaction; PFIN: probabilsitc functional integrated network.

Species	Methodology	Proteins	Interactions	Ref
Arabidopsis lyrata	Integrated	3,979	37,163	[13]
Brachypodium distachyon	Interolog	-	105,705	[238]
	Integrated	3,326	25,063	[13]
Brassica rapa	Interolog, DDI	-	723,310	[258]
Camellia sinensis	Interolog, gene expression	11,208	197,820	[260]
	Interolog, DDI	12,033	216,107	[92]
Chlamydomonas reinhardtii	Interolog	-	49,350	[238]
	Phylogenetic profile	1,086	11,094	[268]
	Integrated	1,449	8,207	[13]
Citrus sinensis	Interolog	8,195	124,491	[263]
Coffea canephora	Interolog	939	4,587	[259]
Glycine max	Interolog	-	160,024	[238]
	Machine learning	-	13,527,834	[231]
	Integrated	5,785	89,538	[13]
Hordeum vulgare	PFIN	26,145	1,272,200	[254]
	Integrated	50	152	[13]
Macrotyloma uniflorum	Interolog, DDI	1,812	6,804	[256]
Manihot esculenta	Interolog, gene expression	24,590	3,638,916	[262]
	Interolog	7,209	90,173	[261]
Medicago truncatula	Interolog	-	112,478	[238]
Ocimum tenuiflorum	Interolog	13,660	327,409	[264]
Oryza sativa	Machine learning	16,895	708,819	[249]
	Machine learning	716	1,341	[188]
	Interolog	-	99,296	[238]
	Interolog	5,049	76,585	[250]
	Interolog, gene expression	-	9,132	[269]
	PFIN	18,377	588,221	[252]
	PFIN	25,765	1,775,000	[253]
	PFIN	4,567	35,441	[248]
	PFIN	12,184	2,996,703	[198]
	Interolog	27,746	14,614,067	[270]
Oryza sativa Indica Group	Integrated	105	224	[13]
Oryza sativa Japonica Group	Integrated	1,787	54,247	[13]
Oryza brachyantha	Integrated	2,946	23,484	[13]
Oryza glaberrima	Integrated	35	100	[13]
Ostreococcus tauri	Integrated	895	3,107	[13]
Ostreococcus lucimarinus	Integrated	1,000	3,855	[13]
Physcomitrella patens	Interolog	5,695	67,740	[271]
	Integrated	3,190	31,790	[13]
Populus trichocarpa	Machine learning	19,321	481,253	[187]
	Interolog	-	135,876	[238]
	Integrated	4,914	60,001	[13]
	Machine learning	1,654	5,536	[188]

Continued on next page

 doi:10.20944/preprints202110.0185.v1

Table 4 – continued from previous page

Species	Methodology	Proteins	Interactions	Ref
Ricinus communis	Interolog	-	99,157	[238]
Selaginella moellendorffii	Interolog	-	112,480	[238]
	Integrated	4,345	52,874	[13]
Setaria italica	Integrated	3,403	27,827	[13]
Solanum lycopersicum	Interolog	10,626	35,7946	[257]
	Interolog	-	110,943	[238]
	Integrated	3,601	32,470	[13]
Solanum tuberosum	Integrated	3,239	27,478	[13]
	Interolog	-	81,057	[238]
Solanum bicolor	Integrated	3,503	26,577	[13]
	Machine learning	751	1,874	[188]
Tetraselmis subcordiformis	Interolog, DDI	938	7,773	[272]
Vitis vinifera	Interolog	-	105,415	[238]
	Integrated	3,418	28,243	[13]
	Machine learning	276	503	[188]
Zea mays	Interolog	6,004	49,026	[255]
	Interolog	-	112,597	[238]
	Machine learning	14,000	2,762,560	[273]
	Machine learning	-	13,175,414	[231]
	Integrated	2,849	18,046	[13]
	PFIN	26,624	922,039	[274]

Protozoa

The apicomplexan *Plasmodium falciparum* is the causative agent of malaria [275]. Interactome analysis is a powerful method for understanding this parasite and identifying potential targets for therapeutic intervention. Few large-scale interaction screens have been carried out in *P. falciparum* due to its AT-rich genome hampering classical experimental methodologies such as yeast two hybrid [276]. A comparison of the *P. falpicarum* interactome [277] with those of yeast, worm, fly and human revealed a marked difference in hub connectivity in the malarial network, with clustered interconnected hubs [278]. This divergence from other species has also been observed in *P. falciparum*'s protein complex conservation [279].

Computational prediction has been applied to host-malaria infection. Dyer and co-workers integrated intra-species PPIs with protein-domain profiles to predict PPIs between *P. falciparum* and its human host [280]. Interolog mapping using eighteen other eukaryotic species was used to produce an interactome for *P. falciparum* and predict its interactions with human proteins [70]. A later interolog study based on multiple interaction datasets revealed parasite proteins predominantly target hub proteins to take control of the human host cell [281]. An integrated interactome of predicted and experimental data was used to study the pathogenesis of cerebral malaria [282] and in the related species, *P. vivax*, machine learning was used to create a human-malaria interactome that was analysed to identify putative drug targets [283].

Leishmania and *Trypanosoma* are groups of protazoan parasites that cause disease in humans—Leishmaniasis and Chagas disease/sleeping sickness, respectively [284–286]—for which there are

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

several interactome datasets. Interactome networks for *Leishmania braziliensis* and *Leishmania infantum* were produced by Dos Santos Vasconce and colleagues based on structural data and machine learning [287]. These predicted networks, were later enhanced by incorporating data from an interolog mapping study in which networks were produced for three Leishmania species, *L. braziliensis*, *L. major* and *L. infantum* before confidence evaluation using gold standard data [288].

To identify potential drug targets for Chagas disease, the predicted secretome for this and the insect pathogen *Trypanosoma rangeli*, were computationally mapped to cellular pathways [289]. Expression data were mapped to interolog networks of *T. brucei* and its vector *Glossina morsitans morsitans* to identify genes and proteins involved in the response to infection [290]. In a later study, interolog mapping identified interactions between *T. brucei* and *G. m. morsitans* [291]. TrypsNetDB contains experimental and interolog interactions for several *Leishmania* and *Trypanosoma* [292].

Several other interactome studies have been carried out in protozoans (Table 5). Date and colleagues created a probabilistic functional integrated network for *P. falciparum* and mapped the network to *T. gondii*, and *C. parvum*, to identify areas of commonality [293]. Twenty three protozoan networks were produced in a study of 68 eukaryotic integrated interactomes [13]. Finally, Cuesta-Astroz and colleague's large-scale study of fifteen parasite-host interolog networks including *Plasmodiums*, *Trypanosomas* , *Leishmanias* and other apicomplexan parasites [294].

Table 5: Computational interactomes in protozoans. DDI: domain-domain interaction; PFIN: probabilisitc functional integrated network. **P. yoelii*, *Toxoplasma gondii*, and *Cryptosporidium parvum*. ***P. falciparum*, *P. knowlesi*, and *P. berghei*.

Species	**Methodology**	**Proteins**	**Interactions**	**Ref**
Cryptosporidium hominis	Integrated	279	590	[13]
Cryptosporidium parvum	Integrated	287	1,153	[13]
	Machine learning	289	487	[188]
Dictyostelium discoideum	Integrated	2,065	24,430	[13]
	Machine learning	1,154	2,458	[188]
Dictyostelium purpureum	Integrated	1,119	4,264	[13]
Entamoeba histolytica	Integrated	641	1,402	[13]
Entamoeba dispar	Integrated	616	1,273	[13]
Giardia lamblia	Integrated	238	453	[13]
Leishmania arabica	Integrated	1,356	5,262	[292]
Leishmania braziliensis	Structural, machine learning	681	6,198	[287]
	Interolog	7,950	39,420	[288]
	Integrated	1,472	6,150	[292]
	Integrated	735	2,490	[13]
	Machine learning	321	621	[188]
Leishmania donovani	Integrated	1,379	5,444	[292]
Leishmania enriettii	Integrated	1,389	5981	[292]
Leishmania infantum	Structural, machine learning	708	7,391	[287]
	Interolog	7,823	45,235	[288]
	Integrated	1,437	6657	[292]
	Integrated	725	2,347	[13]
	Machine learning	316	543	[188]
Leishmania major	Interolog, DDI	1,366	33,861	[295]

Continued on next page

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

Table 5 – continued from previous page

Species	Methodology	Proteins	Interactions	Ref
	Interolog	8,160	43,531	[288]
	Integrated	1,477	7,227	[292]
	Machine learning	367	642	[188]
	Integrated	729	2,537	[13]
Leishmania mexicana	Integrated	1,436	6,597	[292]
Monosiga brevicollis	Integrated	839	2,840	[13]
	Machine learning	621	933	[188]
Naegleria gruberi	Integrated	1,102	4,471	[13]
Plasmodium berghei	Integrated	428	1,056	[13]
Plasmodium chabaudi	Integrated	422	1,035	[13]
Plasmodium falciparum	Interolog	212	344	[296]
	Interolog, gene expression	2,646	286,620	[297]
	PFIN	3,667	388,969	[293]
	PFIN	2,273	133,158	[198]
	Machine learning	1,246	2,551	[188]
	Integrated	1,688	4,634	[13]
	Interolog	-	2,844	[70]
Plasmodium knowlesi	Integrated	444	1,171	[13]
Plasmodium vivax	Interolog	-	14,844	[283]
	Integrated	426	1,212	[13]
Plasmodium yoelii	Integrated	404	1,106	[13]
Paramecium tetraurelia	Integrated	2,082	17,045	[13]
Tetrahymena thermophila	Integrated	1,093	5,323	[13]
Theileria parva	Integrated	334	793	[13]
Theileria annulata	Integrated	365	857	[13]
Toxoplasma gondii	Integrated	596	1,739	[13]
Trichomona vaginalis	Integrated	813	2,611	[13]
Trypanosoma brucei	Integrated	1,685	7,946	[292]
	Integrated	708	3,494	[13]
Trypanosoma brucei gambiense	Integrated	1,644	7,821	[292]
Trypanosoma cruzi	Integrated	1,383	5,728	[292]
	Integrated	1,267	5,258	[13]
Trypanosoma evansi	Integrated	1,704	8,177	[292]
Trypanosoma vivax	Integrated	1,519	6,688	[292]
Multiple species*	PFIN	1,001	47,364	[293]
Multiple species**	Machine learning	1,761	26,060	[298]

Mammals

The laboratory mouse, *Mus musculus*, is possibly the most important model species that has been used extensively in the study of genetics [299]. Due to its importance to the understanding of human disease, the mouse interactome has been a key research goal since completion of the mouse genome, and a wealth of experimental [89, 300–308] and computational [13, 64, 188, 198, 309–312]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

interactome data have been produced, which along with human and yeast data, form the basis of interactome studies in many other mammals (Table 6).

One of the largest, MouseNet, represents a large-scale PFIN produced via machine learning based on multiple types of functional interaction data including PPIs, expression and phenotypes. This network successfully predicts known human disease phenotypes, demonstrating the potential of interactomes in cross-species prediction [64, 313]. In a comparative analysis, Shin and colleagues found that differing ortholog mapping algorithms have low overlap, and so produced an interolog network that combined the different results [310]. Yellaboina and co-workers combined interolog mapping with genome context and phylogenetic profiles to produce a network of over 40,000 mouse interactions [309]. MppDB is a predicted mouse interactome built using text-mined data followed by machine learning using several data types and mapping techniques [311].

Many mammalian studies have concerned species with commercial value, for example *Bos taurus* interactomes were used to investigate meat production [314] and infection [315]. Wang and colleagues predicted interactomes for cattle, dogs, horses and rabbits and demonstrated their reliability using subcellular localisation and in comparison to randomised networks [316]. Of particular note is the evolutionary study of Zitnik and co-workers, who produced interactomes for 1,840 species, including 28 mammals. Interactomes were found to evolve to become more resilient to network failure, and in bacteria this resilience was correlated with the variability of the species' environment [13]. The FunCoup database contains PFINs for sixteen model species including four mammals, and provides an interactive interface for comparative interactomics between species [198]. The STRING database contains functional interaction data, including co-citation, co-expression and gene neighbourhood, for multiple species including several model mammals, which can be queried through an interactive server by protein name [317]. Finally, the BiomeNet server can be used to construct PFINs for any species based a set of eighteen PFINS including human and mouse [318].

Table 6: Computational interactomes in mammals. DDI: domain-domain interaction; PFIN: probabilistic functional integrated networks.

Species	**Methodology**	**Proteins**	**Interactions**	**Ref**
Ailuropoda melanoleuca	Integrated	5,034	41,328	[13]
Bos taurus	Interolog, DDI	17,291	447,014	[316]
	Integrated	8,615	276,128	[13]
	Gene expression	4,995	1,538,522	[319]
	Interolog	330	-	[314]
	Machine learning	7,455	30,049	[188]
	PFIN	17,906	4,551,013	[198]
Callithrix jacchus	Integrated	327	946	[13]
Canis lupus familiaris	Interolog, DDI	13,129	129,386	[316]
	PFIN	17,742	3,853,720	[198]
	Integrated	7,923	166,025	[13]
	Machine learning	4,524	14,478	[188]
Cavia porcellus	Integrated	165	268	[13]
	Machine learning	129	155	[188]
Equus caballus	Interolog, DDI	10,689	93,414	[316]
	Integrated	5,395	45,074	[13]
	Machine learning	205	306	[188]

Continued on next page

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

Table 6 – continued from previous page

Species	Methodology	Proteins	Interactions	Ref
Felis catus	Integrated	5,285	44,325	[13]
Gorilla gorilla	Integrated	5,309	41,774	[13]
Ictidomys tridecemlineatus	Integrated	166	321	[13]
Loxodonta africana	Integrated	171	288	[13]
Macaca mulatta	Integrated	5,003	37,070	[13]
	Machine learning	205	234	[188]
Microcebus murinus	Integrated	64	121	[13]
Monodelphis domestica	Integrated	4,937	39,168	[13]
Myotis lucifugus	Integrated	140	262	[13]
Mustela putorius furo	Integrated	173	254	[13]
Nomascus leucogenys	Integrated	415	715	[13]
Ornithorhynchus anatinus	Integrated	3,817	26,941	[13]
Oryctolagus cuniculus	Interolog, DDI	12,586	115,296	[316]
	Integrated	292	844	[13]
Otolemur garnettii	Integrated	162	428	[13]
Pan troglodytes	Integrated	5,116	39,642	[13]
	Machine learning	195	228	[188]
Pongo abelii	Integrated	5,102	37,676	[13]
	Machine learning	919	1,563	[188]
Procavia capensis	Integrated	57	56	[13]
Pteropus vampyrus	Integrated	107	174	[13]
Rattus norvegicus	Integrated	13,897	305,939	[320]
	Integrated	9,439	261,737	[13]
	Machine learning	6,073	22,300	[188]
	PFIN	18,322	5,560,189	[198]
Sarcophilus harrisii	Integrated	5,004	40,001	[13]
Sus scrofa	Integrated	8,201	143,516	[13]
	Machine learning	2,356	4,323	[188]
	Interolog, DDI	11,955	567,441	[321]
	Interolog	9,534	204,699	[322]
Tursiop truncatus	Integrated	110	163	[13]

Fish

Zebrafish, *Danio rerio*, the best characterised of the fish, is a model for regeneration and development [323], and has been the subject of several experimental interactome studies [13, 188, 198, 324, 325]. Due to the economic importance of global fish production [326], many studies have aimed to understand the interactome of other fish (Table 7) and their responses to parasitic disease [210–212, 327–332].

Carrera and colleagues created an interactome from the STRING database based on the combined proteome of fifteen different sarcoplasmic fish, revealing a core interactome involved in energy and metabolism [333]. Millan-Cubillo and co-workers also used STRING to produce interaction networks for two developmental stages in the seabream, *Sparus aurata*, based on expression data [334]. A later

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

seabream study used expression data to mine STRING for interactions of the stress response [335]. Co-expression analysis was also used to create a genetic interactome for the Nile tilapia, *Oreochromis niloticus* [336]. Zitnik and co-workers evolutionary study included seven fish and the coelocanth, *Latimeria chalumnae* [13].

Many interactome-based studies have been applied to understand the mechanisms of infection in fish. For instance, four related studies created host-pathogen interactomes between the parasites and fish species to investigate different aspects of infection: immune responses in gill tissues of the grouper [329]; interactions with host plasma proteins in carp [328]; interactions with liver cells in grouper [330]; and the blood immune response of the Japanese flounder [332]. Kumar and co-workers studied the immune response to *Yersinia ruckeri* infection in the rainbow trout, *Oncorhynchus mykiss*, using a combination of expression data and the STRING database [331]. Several studies have also investigated *C. albicans* infection in zebrafish using microarray data [210] and interolog mapping [211, 212] to decipher the zerbrafish immune response.

Table 7: Computational interactomes in Fish. **Gadus morhua*, *Merluccius australis australis*, *Lophius piscatorius*, *Genypterus blacodes*,*Brama brama*,*Diplodus sargus*,*Pagellus bogaraveo*,*Scomber japonicus*,*Sparus aurata*, *Thunnus albacares*, *Trachurus trachurus*, *Xiphias gladius*, *Lepidorhombus boscii*, *Solea solea*, *Salmo salar*.

Species	Methodology	Proteins	Interactions	Reference
Gasterosteus aculeatus	Integrated	45	57	[13]
Latimeria chalumnae	Integrated	3,826	27,789	[13]
Oreochromis niloticus	Integrated	32	55	[13]
Oryzias latipes	Integrated	4,134	35,414	[13]
Sparus aurata	Integrated	18	61	[335]
Takifugu rubripes	Integrated	4,111	35,801	[13]
Tetraodon nigroviridis	Integrated	41	59	[13]
	Machine learning	142	203	[188]
Xiphophorus maculatus	Integrated	4,503	41,341	[13]
Multiple species*	Integrated	84	279	[333]

Insects

Drosphila melanogaster, the fruit fly, is the model insect and is widely used in genetic studies [337]. Several studies have been dedicated to understanding the interactome of this important species [13, 188, 198, 307, 338–355]. The Drosophila Interactions Database (DroID) [355, 356] and the Predicted Drosophila Interactome Resource (PDIR) are integrated database containing both experimental and predicted interactions [354]. Interactome analyses have been applied to the disease vectors *Glossina morsitans morsitans* (tsetse flies) [290, 291], *Aedes aegypti* [357–359] and *Anopheles gambiae* [13, 188] (mosquitos), and the pests *Acyrthosiphon pisum* (aphid) and *Tribolium castaneum* (flour beetle) [13]. The interactomes of several other insect species, such as the carpenter ant *Camponotus floridanus* [360], and the silk moth *Bombyx mori* [361] have also been investigated (Table 8).

Nematodes and Platyhelminths

The vast majority of experimental interactions for nematodes are for the model worm *Caenorhabditis elegans* [307, 353, 362–373]. *C. elegans* is a free-living nematode that was one of the first eukaryotes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

Table 8: Computational interactomes in Insects. DDI: domain-domain interactions; PFIN: probabilistic functional interaction network.

Species	Methodology	Proteins	Interactions	Reference
Acyrthosiphon pisum	Integrated	1,867	8,354	[13]
Aedes aegypti	PFIN	4,214	10,209	[357]
	Machine learning	1,310	2,396	[188]
	Integrated	1,994	8,913	[13]
Anopheles gambiae	Machine learning	1,537	3,442	[188]
	Integrated	1,808	7,214	[13]
Apis mellifera	Integrated	1,561	6,583	[13]
Bombyx mori	Interolog	4,623	7,736	[361]
	Integrated	1,732	8,512	[13]
Camponotus floridanus	Interolog, DDI	6,274	51,866	[360]
Culex quinquefasciatus	Integrated	1,846	8,800	[13]
Drosophila ananassae	Integrated	2,078	11,429	[13]
Drosophila grimshawi	Integrated	2,203	13,143	[13]
Drosophila pseudoobscura	Machine learning	1,461	2,650	[188]
	Integrated	2,111	11,062	[13]
Drosophila virilis	Machine learning	2,163	5,072	[188]
	Integrated	2,144	11,561	[13]
Drosophila yakuba	Integrated	2,119	11,268	[13]
Drosophila willistoni	Integrated	2,092	10,739	[13]
	Machine learning	2,175	5,261	[188]
Nasonia vitripennis	Integrated	1,397	4,805	[13]
Pediculus humanus	Integrated	1,710	7,539	[13]
	Machine learning	1,043	1,812	[188]
Tribolium castaneum	Integrated	1,734	8,073	[13]

to have its genome sequenced [374], and considerable effort has been put into completing the interactome of this species [375], and in the comparison between its interactome and those of other model eukaryotes [307]. Several computationally-predicted interactomes have also been produced [13, 188, 198, 376–381].

Li and colleagues combined physical interaction data with interolog mapping to produce the Worm Interactome [376]. Simonis and co-workers extended this network using further Y2H screening to produce version 8 of the Worm Interactome [377]. By combining Y2H and protein–DNA interaction (PDI) mapping the *C. elegans* interactome was expanded to include more than 2000 transcription factor interactions [378]. Gunsalus and co-workers combined the Worm Interactome (version 5) [376] with expression and phenotypic data to produce an integrated network of early embryogenesis [379]. These datasets have formed the basis of many interactome analyses and comparisons in *C. elegans* and beyond [382, 383].

Few interactome studies have been carried out in other nematodes (Table 9) Interolog mapping was used to produce and compare host parasite interactomes for six parasites including the human and plant parasites *Meloidogyne hapla* and *Meloidogyne incognita* [384]. Comparison with a predicted interactome for *C. elegans* was then used to prioritise drug targets. Finally, Cuesta-Astroz and colleague's large-scale study of fifteen parasite-host interaction networks included an interactome network for *Trichinella spiralis* [294].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

The platyhelminths *Schistosoma mansoni* and *Schistosoma japonicum* are important parasitic blood flukes that cause schistosomiasis in humans [385, 386]. Interolog mapping has been used to produce and compare host-parasite interactomes for six parasites including *S. mansoni* and *S. japonicum* [384]. White-Bear and colleagues used structural prediction, followed by extensive confidence filtering, to produce an interactome of over 1000 *S. mansoni*-human interactions [387]. A combination of Y2H and Co-IP produced 205 interactions involving the essential histone deacetylase 8 [388]. Co-IP was also used to identify the host-parasite interactome of *S. mansoni* and its mollusc host *Biomphalaria glabrata* [389]. In Cuesta-Astroz and colleague's large-scale study of fifteen parasite-host interaction networks, a network of $\sim$ 700 interactions was produced for *S. mansoni* [294].

Castillo-Lara and co-workers used interolog mapping from a human reference interactome, gene expression data and machine learning to produce PlanNet, a predicted interactome for the planarian, *Schmidtea mediterranea*, with online visualisation and analysis tool [390]. This resource was later extended to allow exploration of the network using gene expression data [391]. Finally, a probabilisitic functional integrated network of interologs was produced for the mouse bile duct tapeworm, *Hymenolepis microstoma* [74]. Although there is little interactome data for these species (Table 9), interactomics has the potential to expand our understanding of parasitism in the future.

Table 9: Computational interactomes in Nematodes and Platyhelminths. PFIN: probabilistic functional interaction network. * *Brugia malayi*, *Meloidogyne hapla*, *Meloidogyne incognita*, *Trichinella spiralis*, *Schistosoma mansoni* and *Schistosoma japonicum*; networks combined to predict common drug target PPIs.

Species	Methodology	Proteins	Interactions	Reference
Caenorhabditis briggsae	Machine learning	474	699	[188]
	Integrated	1,547	6,558	[13]
Caenorhabditis remanei	Machine learning	1,312	2,225	[188]
Trichinella spiralis	Integrated	1,453	5,511	[13]
Hymenolepi microstoma	PFIN	3,474	20,684	[74]
Schistosoma mansoni	Integrated	958	3,520	[13]
Schmidtea mediterranea	Interolog, machine learning	-	729,043	[390]
Multiple*	Interolog	-	8	[384]

Other Eukaryotes

Interactome analysis that have applied to a variety of other species, including birds, amphibians, reptiles and crustaceans, are summarised in Table 10.

Table 10: Computational interactomes in other eukaryotes. DDI: domain-domain interaction; PFIN: probabilistic functional interaction network.

Species	Methodology	Proteins	Interactions	Ref
Birds				
Gallus gallus	Interolog, DDI	12,207	328,590	[316]
	Interolog	8,140	72,000	[392]
	Integrated	5,208	141886	[13]

Continued on next page

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 October 2021 doi:10.20944/preprints202110.0185.v1

Table 10 – continued from previous page

Species	Methodology	Proteins	Interactions	Ref
	Machine learning	4,542	14,244	[188]
	PFIN	12,289	1,608,939	[198]
Meleagris gallopavo	Integrated	2,656	30,024	[13]
Taeniopygia guttata	Integrated	6,105	227,682	[13]
Amphibians				
Xenopus tropicalis	Integrated	6,509	130,423	[13]
	Machine learning	966	1,887	[188]
Reptiles				
Pelodiscus sinensis	Integrated	3,464	24,446	[13]
Anolis carolinensis	Integrated	3,316	23,485	[13]
Molluscs				
Theba pisana	DDI	3,913	41,653	[393]
Crustaceans				
Daphnia pulex	Machine leanring	1,122	1,977	[188]
Eriocheir sinensis	Interolog	3,223	35,787	[394]
Litopenaeus vannamei	Interolog	4,858	104,187	[395]
	Gene expression, interolog	127	1292	[396]
	Gene expression, interolog	383	8,557	[397]
Arachnids				
Stegodyphus mimosarum	Interolog	3,810	58,489	[398]
Ixodes scapularis	Integrated	2,052	12,610	[13]
	Machine learning	749	1,103	[188]
Echinodermata				
Strongylocentrotus purpuratus	Integrated	1,746	8,732	[13]
Cnidaria				
Nematostella vectensis	Machine learning	1,135	2,842	[188]
	Integrated	1,621	7,487	[13]
Heterokonta				
Phaeodactylum tricornutum	Integrated	980	4,057	[13]
Phytophthora infestans	Integrated	1,507	8,803	[13]
Thalassiosira pseudonana	Integrated	1,053	4,739	[13]
Rhodophyta				
Cyanidioschyzon merolae	Integrated	895	3,657	[13]
Placozoa				
Trichoplax adhaerens	Integrated	1,353	5,610	[13]
	Machine learning	1,022	1,688	[188]
Porifera				
Amphimedon queenslandica	Integrated	1,487	6,853	[13]
Cephalochordata				
Branchiostoma floridae	Integrated	3,015	31,607	[13]
	Machine learning	1,481	4,143	[188]
Tunicata				
Ciona intestinalis	Integrated	1,372	5,432	[13]
	PFIN	6,098	1,373,106	[198]
Ostreococcus lucimarinus	Machine learning	576	946	[188]

Continued on next page

 doi:10.20944/preprints202110.0185.v1

Table 10 – continued from previous page

Species	Methodology	Proteins	Interactions	Ref

Future Perspective

Data from traditional model organisms, chosen in part for ease of their experimental study [399], are now being expanded by the addition of genomic data from closely-related species [400]. Diverse organisms have become the study species of choice for answering more obscure biological questions [401] and to test the long-held fundamental rules built from the original model species [402]. Comparative interactomics is now possible on a large scale as demonstrated by Zitnik and colleagues study of 1,840 predicted interactomes spanning the tree of life [13]. With over 8 million protein-protein interactions, this dataset allowed the authors to observe interaction rewiring that had only previously been seen on a smaller scale [191–193, 353, 378, 403].

Computationally-predicted interactions can compensate for the lack of experimental data in many species but, like experimental methods, they have drawbacks such as noise and false positive interactions [293, 404]. Conservation of PPI is unequal and accuracy of interolog mapping will vary between species [405]. This accuracy will also be dependent on evolutionary distance and careful selection of thresholds will be necessary [406]. It's clear from yeast interactomes that there can be low overlap between species; *S. cerevsiae* is well characterised but the *C. albicans* and *S. pombe* interactomes are different in a number of respects [191, 193, 204, 407]. Combining different computational methods [86, 105, 106, 310], and experimental data if available [232, 235, 376, 408], can give a fuller picture of the interactome and mitigate the effect of data noise by using a probabilistic framework [74, 198].

Interolog and DDI mapping can only detect interactions within conserved sequences or domains [68]; organism-specific proteins and interactions are missed. Filling in the gaps is vital to our understanding of biology and evolution. Interactome analyses can help to identify these gaps and target further analysis in non-model species. While there are parts of some interactomes that cannot currently be predicted, interactome accuracy will improve as coverage of diverse species increases. Non-model species have already been used to provide insights in a number of fields, including human disease [409], ageing and regeneration [410, 411], and to expand our understanding of evolution [412], ecology [413, 414] and biological diversity [415, 416]. There are currently more than 13,000 entries from over 10,000 species in the NCBI Genome database[1], the majority of which are non-models. At the same time, *de novo* transcriptomes have also been produced for many species that lack genome sequence data [417–420]. The Sequence Read archive contains more than 20,000 non-model transcriptomic datasets from over 7,000 species[1]. These resources represent a huge amount of data for interactome generation in more diverse species and provide the basis to build interactomes that will potentially have impact across the field of Biology.

References

[1] A H Tong, M Evangelista, A B Parsons, H Xu, et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294 (2001), pp. 2364–2368.

[2] X Pan, D S Yuan, D Xiang, X Wang, et al. A robust toolkit for functional profiling of the yeast genome. *Mol Cell* 16 (2004), pp. 487–496.

[3] M Schuldiner, S R Collins, N J Thompson, V Denic, et al. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* 123 (2005), pp. 507–519.

[1] January 2021

[4] T Ito, T Chiba, R Ozawa, M Yoshida, et al. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc Natl Acad Sci U S A* 98 (2001), pp. 4569–4574.

[5] P Uetz, L Giot, G Cagney, T A Mansfield, et al. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403 (2000), pp. 623–627.

[6] A C Gavin, P Aloy, P Grandi, R Krause, et al. Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440 (2006), pp. 631–636.

[7] N J Krogan, G Cagney, H Yu, G Zhong, et al. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440 (2006), pp. 637–643.

[8] R Yu and J Nielsen. Big data in yeast systems biology. *FEMS Yeast Res* 19 (2019), foz070.

[9] L Loewe. A framework for evolutionary systems biology. *BMC Syst Biol* 3 (2009), p. 27.

[10] T Ideker, T Galitski, and L Hood. A new approach to decoding life: systems biology. *Annu Rev Genomics Hum Genet* 2 (2001), pp. 343–372.

[11] A-L Barabási and Z N Oltvai. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5 (2004), pp. 101–113.

[12] E V Koonin and Y I Wolf. Evolutionary systems biology: links between gene evolution and function. *Curr Opin Biotechnol* 17 (2006), pp. 481–487.

[13] M Zitnik, R Sosič, M W Feldman, and J Leskovec. Evolution of resilience in protein interactomes across the tree of life. *Proc Natl Acad Sci U S A* 116 (2019), pp. 4426–4433.

[14] T Lemberger. Systems biology in human health and disease. *Mol Syst Biol* 3 (2007), p. 136.

[15] H Kitano. Systems biology: a brief overview. *Science* 295 (2002), pp. 1662–1664.

[16] M E Cusick, N Klitgord, M Vidal, and D E Hill. Interactome: gateway into systems biology. *Hum Mol Genet* 14 Spec No. 2 (2005), pp. 171–181.

[17] J Goll and P Uetz. The elusive yeast interactome. *Genome Biol* 7 (2006), p. 223.

[18] S H Yook, Z N Oltvai, and A-L Barabási. Functional and topological characterization of protein interaction networks. *Proteomics* 4 (2004), pp. 928–942.

[19] L H Hartwell, J J Hopfield, S Leibler, and A W Murray. From molecular to modular cell biology. *Nature* 402 (1999), pp. C47–C52.

[20] D Eisenberg, E M Marcotte, I. Xenarios, and T. O. Yeates. Protein function in the post-genomic era. *Nature* 405 (2000), pp. 823–826.

[21] M Brehme and M Vidal. A global protein-lipid interactome map. *Mol Syst Biol* 6 (2010), p. 443.

[22] A Trewavas. A brief history of systems biology. *Plant Cell* 18 (2006), pp. 2420–2430.

[23] I Lee, S V Date, A T Adai, and E M Marcotte. A probabilistic functional network of yeast genes. *Science* 306 (2004), pp. 1555–1558.

[24] I Lee, Z Li, and E M Marcotte. An improved, bias-reduced probabilistic functional gene network of baker's yeast, *Saccharomyces cerevisiae*. *PloS One* 2 (2007), e988.

[25] K James, A Wipat, and J Hallinan. Integration of full-coverage probabilistic functional networks with relevance to specific biological processes. *Data Integr Life Sci* 5647 (2009), pp. 31–46.

[26] W Huber, V J Carey, L Long, S Falcon, et al. Graphs in molecular biology. *BMC Bioinformatics* 8 (2007), S8.

 doi:10.20944/preprints202110.0185.v1

[27] G C K W Koh, P Porras, B Aranda, H Hermjakob, et al. Analyzing protein-protein interaction networks. *J Proteome Res* 11 (2012), pp. 2014–2031.

[28] O Mason and M Verwoerd. Graph theory and networks in Biology. *IET Syst Biol* 1 (2007), pp. 89–119.

[29] D Bu, Y Zhao, L Cai, H Xue, et al. Topological structure analysis of the protein-protein interaction network in budding yeast. *Nucleic Acids Res* 31 (2003), pp. 2443–2450.

[30] H Jeong, S P Mason, A L Barabási, and Z N Oltvai. Lethality and centrality in protein networks. *Nature* 411 (2001), pp. 41–42.

[31] S Wuchty, Z N Oltvai, and A L Barabási. Evolutionary conservation of motif constituents in the yeast protein interaction network. *Nat Genet* 35 (2003), pp. 176–179.

[32] S Asthana, O D King, F D Gibbons, and F P Roth. Predicting protein complex membership using probabilistic network reliability. *Genome Res* 14 (2004), pp. 1170–1175.

[33] G T Hart, I Lee, and E R Marcotte. A high-accuracy consensus map of yeast protein complexes reveals modular nature of gene essentiality. *BMC Bioinformatics* 8 (2007), p. 236.

[34] L Ou-Yang, X-F Zhang, D-Q Dai, M-Y Wu, et al. Protein complex detection based on partially shared multi-view clustering. *BMC Bioinformatics* 17 (2016), p. 371.

[35] D Piovesan, M Giollo, E Leonardi, C Ferrari, et al. INGA: protein function prediction combining interaction networks, domain assignments and sequence similarity. *Nucleic Acids Res* 43 (2015), W134–W140.

[36] H N Chua, W K Sung, and L Wong. Using indirect protein interactions for the prediction of Gene Ontology functions. *BMC Bioinformatics* 8 (2007), S8.

[37] U Karaoz, T M Murali, S Letovsky, Y Zheng, et al. Whole-genome annotation by using evidence integration in functional-linkage networks. *Proc Natl Acad Sci U S A* 101 (2004), pp. 2888–2893.

[38] C S Goh and F E Cohen. Co-evolutionary analysis reveals insights into protein-protein interactions. *J Mol Biol* 324 (2002), pp. 177–192.

[39] H B Fraser, A E Hirsh, L M Steinmetz, C Scharfe, et al. Evolutionary rate in the protein interaction network. *Science* 296 (2002), pp. 750–752.

[40] B P Kelley, R Sharan, R M Karp, T Sittler, et al. Conserved pathways within bacteria and yeast as revealed by global protein network alignment. *Proc Natl Acad Sci U S A* 100 (2003), pp. 11394–11399.

[41] M Jafari, M Sadeghi, M Mirzaie, S-A Marashi, et al. Evolutionarily conserved motifs and modules in mitochondrial protein-protein interaction networks. *Mitochondrion* 13 (2013), pp. 668–675.

[42] A Clauset, C Moore, and M E Newman. Hierarchical structure and the prediction of missing links in networks. *Nature* 453 (2008), pp. 98–101.

[43] I A Kovács, K Luck, K Spirohn, Y Wang, et al. Network-based prediction of protein interactions. *Nat Commun* 10 (2019), p. 1240.

[44] E Franzosa, B Linghu, and Y Xia. Computational reconstruction of protein-protein interaction networks: algorithms and issues. *Methods Mol Biol* 541 (2009), pp. 89–100.

[45] M Fromont-Racine, J C Rain, and P Legrain. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* 16 (1997), pp. 277–282.

[46] A C Gavin, M Bosche, R Krause, P Grandi, et al. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415 (2002), pp. 141–147.

[47] Y Ho, A Gruhler, A Heilbut, G D Bader, et al. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415 (2002), pp. 180–183.

[48] J F Rual, K Venkatesan, T Hao, T Hirozane-Kishikawa, et al. Towards a proteome-scale map of the human protein-protein interaction network. *Nature* 437 (2005), pp. 1173–1178.

[49] U Stelzl, U Worm, M Lalowski, C Haenig, et al. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 122 (2005), pp. 957–968.

[50] R M Ewing, P Chu, F Elisma, H Li, et al. Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol Syst Biol* 3 (2007), p. 89.

[51] H Yu, P Braun, M A Yildirim, I Lemmens, et al. High-quality binary protein interaction map of the yeast interactome network. *Science* 322 (2008), pp. 104–110.

[52] C Lee. Coimmunoprecipitation assay. *Methods Mol Biol* 362 (2007), pp. 401–406.

[53] S Pang C N Iand Ballouz, D Weissberger, L M Thibaut, J J Hamey, et al. Analytical guidelines for co-fractionation mass spectrometry obtained through global profiling of gold standard *Saccharomyces cerevisiae* protein complexes. *Mol Cell Proteomics* 19 (2020), pp. 1876–1895.

[54] X Tang and J E. Bruce. Chemical cross-linking for protein-protein interaction studies. *Methods Mol Biol* 492 (2009), pp. 283–293.

[55] S Mohr, C Bakal, and N Perrimon. Genomic screening with RNAi: results and challenges. *Annu Rev Biochem* 79 (2010), pp. 37–64.

[56] A Beyer, S Bandyopadhyay, and T Ideker. Integrating physical and genetic maps: from genomes to interaction networks. *Nat Rev Genet* 8 (2007), pp. 699–710.

[57] S J Dixon, Y Fedyshyn, J L Y Koh, T S K Prasad, et al. Significant conservation of synthetic lethal genetic interaction networks between distantly related eukaryotes. *Proc Natl Acad Sci U S A* 105 (2008), pp. 16653–16658.

[58] C von Mering, E M Zdobnov, S Tsoka, F D Ciccarelli, et al. Genome evolution reveals biochemical networks and functional modules. *Proc Natl Acad Sci U S A* 100 (2003), pp. 15428–15433.

[59] R Sharan and T Ideker. Modeling cellular machinery through biological network comparison. *Nat Biotechnol* 24 (2006), pp. 427–433.

[60] D Banky, R Ordog, and V Grolmusz. NASCENT: an automatic protein interaction network generation tool for non-model organisms. *Bioinformation* 3 (2009), pp. 361–363.

[61] F He, Y Zhang, H Chen, Z Zhang, et al. The prediction of protein-protein interaction networks in rice blast fungus. *BMC Genomics* 9 (2008), p. 519.

[62] X He and J Zhang. Higher duplicability of less important genes in yeast genomes. *Mol Biol Evol* 23 (2006), pp. 144–151.

[63] T A Boltz, P Devkota, and S Wuchty. Collective influencers in protein interaction networks. *Sci Rep* 9 (2019), p. 3948.

[64] E Kim, S Hwang, H Kim, H Shim, et al. MouseNet v2: a database of gene networks for studying the laboratory mouse and eight other model vertebrates. *Nucleic Acids Res* 44 (2016), pp. D848–D854.

[65] S Wuchty, A-L Barabási, and M T Ferdig. Stable evolutionary signal in a yeast protein interaction network. *BMC Evol Biol* 6 (2006), p. 8.

[66] A N Nguyen Ba, B J Yeh, D van Dyk, A R Davidson, et al. Proteome-wide discovery of evolutionary conserved sequences in disordered regions. *Sci Signal* 5 (2012), rs1.

[67] B Snel, V van Noort, and M A Huynen. Gene co-regulation is highly conserved in the evolution of eukaryotes and prokaryotes. *Nucleic Acids Res* 32 (2004), pp. 4725–4731.

[68] L R Matthews, P Vaglio, J Reboul, H Ge, et al. Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or "interologs". *Genome Res* 11 (2001), pp. 2120–2126.

[69] H Yu, N M Luscombe, H X Lu, X Zhu, et al. Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome Res* 14 (2004), pp. 1107–1118.

[70] S-A Lee, C-H Chan, C-H Tsai, J-M Lai, et al. Ortholog-based protein-protein interaction prediction and its application to inter-species interactions. *BMC Bioinformatics* 9 Suppl 12 (2008), S11.

[71] M Michaut, S Kerrien, L Montecchi-Palazzi, F Chauvat, et al. InteroPorc: automated inference of highly conserved protein interaction networks. *Bioinformatics* 24 (2008), pp. 1625–1631.

[72] T W Huang, A C Tien, W S Huang, Y C Lee, et al. POINT: a database for the prediction of protein-protein interactions based on the orthologous interactome. *Bioinformatics* 20 (2004), pp. 3273–3276.

[73] A M Wiles, M Doderer, J Ruan, T-T Gu, et al. Building and analyzing protein interactome networks by cross-species comparisons. *BMC Syst Biol* 4 (2010), p. 36.

[74] K James and P D Olson. The tapeworm interactome: inferring confidence scored protein-protein interactions from the proteome of *Hymenolepis microstoma*. *BMC Genomics* 21 (2020), p. 346.

[75] J M Stuart, E Segal, D Koller, and S K Kim. A gene-coexpression network for global discovery of conserved genetic modules. *Science* 302 (2003), pp. 249–255.

[76] P Tsaparas, L Mariño-Ramírez, O Bodenreider, E V Koonin, et al. Global similarity and local divergence in human and mouse gene co-expression networks. *BMC Evol Biol* 6 (2006), p. 70.

[77] V van Noort, B Snel, and M A Huynen. Predicting gene function by conserved co-expression. *Trends Genet* 19 (2003), pp. 238–242.

[78] J Park, M Lappe, and S A Teichmann. Mapping protein family interactions: intramolecular and intermolecular protein family interaction repertoires in the PDB and yeast. *J Mol Biol* 307 (2001), pp. 929–938.

[79] M Deng, S Mehta, F Sun, and T Chen. Inferring domain-domain interactions from protein-protein interactions. *Genome Res* 12 (2002), pp. 1540–1548.

[80] T Doerks, R R Copley, J Schultz, C P Ponting, et al. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res* 12 (2002), pp. 47–56.

[81] E Sprinzak and H Margalit. Correlated sequence-signatures as markers of protein-protein interaction. *J Mol Biol* 311 (2001), pp. 681–692.

[82] S-K Ng, Z Zhang, and S-H Tan. Integrative approach for computationally inferring protein domain interactions. *Bioinformatics* 19 (2003), pp. 923–929.

 doi:10.20944/preprints202110.0185.v1

[83] X W Chen, M Liu, and R Ward. Protein function assignment through mining cross-species protein-protein interactions. *PLoS One* 3 (2008), e1562.

[84] C Chen, J-F Zhao, Q Huang, R-S Wang, et al. Inferring domain-domain interactions from protein-protein interactions in the complex network conformation. *BMC Syst Biol* 6 Suppl 1 (2012), S7.

[85] J R Bock and D A Gough. Predicting protein–protein interactions from primary structure. *Bioinformatics* 17 (2001), pp. 455–460.

[86] D Betel, K E Breitkreuz, R Isserlin, D Dewar-Darch, et al. Structure-templated predictions of novel protein interactions from sequence information. *PLoS Comput Biol* 3 (2007), pp. 1783–1789.

[87] R D Finn, B L Miller, J Clements, and A Bateman. iPfam: a database of protein family and domain interactions found in the Protein Data Bank. *Nucleic Acids Res* 42 (2014), pp. D364–D373.

[88] X-M Zhao, X-W Zhang, W-H Tang, and L Chen. FPPI: *Fusarium graminearum* protein-protein interaction database. *J Proteome Res* 8 (2009), pp. 4714–4721.

[89] K Zhang, Y Li, T Li, Z-G Li, et al. Pathogenicity genes in *Ustilaginoidea virens* revealed by a predicted protein-protein interaction network. *J Proteome Res* 16 (2017), pp. 1193–1206.

[90] T Huo, W Liu, Y Guo, C Yang, et al. Prediction of host - pathogen protein interactions between *Mycobacterium tuberculosis* and *Homo sapiens* using sequence motifs. *BMC Bioinformatics* 16 (2015), p. 100.

[91] D Lei, R Lin, P Yin Cand Li, and A Zheng. Global protein-protein interaction network of rice sheath blight pathogen. *J Proteome Res* 13 (2014), pp. 3277–3293.

[92] G Singh, V Singh, and V Singh. Genome-wide interologous interactome map (TeaGPIN) of *Camellia sinensis*. *Genomics* S0888-7543 (2020), pp. 30452–30455.

[93] O Dror, D Schneidman-Duhovny, A Shulman-Peleg, R Nussinov, et al. Structural similarity of genetically interacting proteins. *BMC Syst Biol* 2 (2008), p. 69.

[94] P Aloy and R B Russell. Interrogating protein interaction networks through structural biology. *Proc Natl Acad Sci U S A* 99 (2002), pp. 5896–5901.

[95] J Janin, K Henrick, J Moult, L T Eyck, et al. CAPRI: a Critical Assessment of PRedicted Interactions. *Proteins* 52 (2003), pp. 2–9.

[96] L Lu, A K Arakaki, H Lu, and J Skolnick. Multimeric threading-based prediction of protein-protein interactions on a genomic scale: application to the *Saccharomyces cerevisiae* proteome. *Genome Res* 13 (2003), pp. 1146–1154.

[97] P Aloy and R B Russell. InterPreTS: protein interaction prediction through tertiary structure. *Bioinformatics* 19 (2003), pp. 161–162.

[98] M Hue, M Riffle, J P Vert, and W S Noble. Large-scale prediction of protein-protein interactions from structures. *BMC Bioinformatics* 11 (2010), p. 144.

[99] L Lu, H Lu, and J Skolnick. MULTIPROSPECTOR: an algorithm for the prediction of protein-protein interactions by multimeric threading. *Proteins* 49 (2002), pp. 350–364.

[100] G R Smith and M J Sternberg. Prediction of protein-protein interactions by docking methods. *Curr Opin Struct Biol* 12 (2002), pp. 28–35.

[101] U Ogmen, O Keskin, A S Aytuna, R Nussinov, et al. PRISM: protein interactions by structural matching. *Nucleic Acids Res* 33 (2005), W331–W336.

[102] P Aloy, B Böttcher, H Ceulemans, C Leutwein, et al. Structure-based assembly of protein complexes in yeast. *Science* 303 (2004), pp. 2026–2029.

[103] M Ohue, Y Matsuzaki, T Shimoda, T Ishida, et al. Highly precise protein-protein interaction prediction based on consensus between template-based and *de novo* docking methods. *BMC Proc* 7 (2013), S6.

[104] Y-S Lo, Y-C Chen, and J-M Yang. 3D-interologs: an evolution database of physical protein-protein interactions across multiple genomes. *BMC Genomics* 11 Suppl 3 (2010), S7.

[105] A S Aytuna, A Gursoy, and O Keskin. Prediction of protein-protein interactions by combining structure and sequence conservation in protein interfaces. *Bioinformatics* 21 (2005), pp. 2850–2855.

[106] S Gong, G Yoon, I Jang, D Bolser, et al. PSIbase: a database of protein structural interactome map (PSIMAP). *Bioinformatics* 21 (2005), pp. 2541–2543.

[107] F Pazos and A Valencia. *In silico* two-hybrid system for the selection of physically interacting protein pairs. *Proteins* 47 (2002), pp. 219–227.

[108] L Skrabanek, H K Saini, G D Bader, and A J Enright. Computational prediction of protein-protein interactions. *Mol Biotechnol* 38 (2008), pp. 1–17.

[109] T Dandekar, B Snel, M Huynen, and P Bork. Conservation of gene order: a fingerprint of proteins that physically interact. *Trends Biochem Sci* 23 (1998), pp. 324–328.

[110] P M Bowers, M Pellegrini, M J Thompson, J Fierro, et al. Prolinks: a database of protein functional linkages derived from coevolution. *Genome Biol* 5 (2004), R35.

[111] R Overbeek, M Fonstein, M D'Souza, G D Pusch, et al. The use of gene clusters to infer functional coupling. *Proc Natl Acad Sci U S A* 96 (1999), pp. 2896–2901.

[112] M Huynen, B Snel, W Lathe, and P Bork. Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. *Genome Res* 10 (2000), pp. 1204–1210.

[113] I K Kirk, N Weinhold, S Brunak, and K Belling. The impact of the protein interactome on the syntenic structure of mammalian genomes. *PloS One* 12 (2017), e0179112.

[114] E M Marcotte, M Pellegrini, H L Ng, D W Rice, et al. Detecting protein function and protein-protein interactions from genome sequences. *Science* 285 (1999), pp. 751–753.

[115] A J Enright, I Iliopoulos, N C Kyrpides, and C A Ouzounis. Protein interaction maps for complete genomes based on gene fusion events. *Nature* 402 (1999), pp. 86–90.

[116] I Yanai, A Derti, and C DeLisi. 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 (2001), pp. 7940–7945.

[117] F Pazos, J A Ranea, D Juan, and M J Sternberg. Assessing protein co-evolution in the context of the tree of life assists in the prediction of the interactome. *J Mol Biol* 352 (2005), pp. 1002–1015.

[118] J Wu, S Kasif, and C DeLisi. Identification of functional links between genes using phylogenetic profiles. *Bioinformatics* 19 (2003), pp. 1524–1530.

[119] Y I Wolf, I B Rogozin, A S Kondrashov, and E V Koonin. Genome alignment, evolution of prokaryotic genome organization, and prediction of gene function using genomic context. *Genome Res* 11 (2001), pp. 356–372.

[120] J Sun, Y Li, and Z Zhao. Phylogenetic profiles for the prediction of protein-protein interactions: how to select reference organisms? *Biochem Biophys Res Commun* 353 (2007), pp. 985–991.

[121] J Sun and Z Zhao. Construction of phylogenetic profiles based on the genetic distance of hundreds of genomes. *Biochem Biophys Res Commun* 355 (2007), pp. 849–853.

[122] R Jothi, M G Kann, and T M Przytycka. Predicting protein-protein interaction by searching evolutionary tree automorphism space. *Bioinformatics* 21 Suppl 1 (2005), pp. i241–i250.

[123] T Gaasterland and M A Ragan. Microbial genescapes: phyletic and functional patterns of ORF distribution among prokaryotes. *Microb Comp Genomics* 3 (1998), pp. 199–217.

[124] M Pellegrini, E M Marcotte, M J Thompson, D Eisenberg, et al. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc Natl Acad Sci U S A* 96 (1999), pp. 4285–4288.

[125] F Pazos and A Valencia. Similarity of phylogenetic trees as indicator of protein-protein interaction. *Protein Eng* 14 (2001), pp. 609–614.

[126] D Juan, F Pazos, and A Valencia. High-confidence prediction of global interactomes based on genome-wide coevolutionary networks. *Proc Natl Acad Sci U S A* 105 (2008), pp. 934–939.

[127] T Sato, Y Yamanishi, M Kanehisa, and H Toh. The inference of protein-protein interactions by co-evolutionary analysis is improved by excluding the information about the phylogenetic relationships. *Bioinformatics* 21 (2005), pp. 3482–3489.

[128] G D Bader and C W Hogue. Analyzing yeast protein-protein interaction data obtained from different sources. *Nat Biotechnol* 20 (2002), pp. 991–997.

[129] S R Collins, P Kemmeren, X C Zhao, J F Greenblatt, et al. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Mol Cell Proteomics* 6 (2007), pp. 439–450.

[130] M E Futschik, G Chaurasia, and H Herzel. Comparison of human protein-protein interaction maps. *Bioinformatics* 23 (2007), pp. 605–611.

[131] A H Tong, G Lesage, G D Bader, H Ding, et al. Global mapping of the yeast genetic interaction network. *Science* 303 (2004), pp. 808–813.

[132] C Huttenhower and O G Troyanskaya. Assessing the functional structure of genomic data. *Bioinformatics* 24 (2008), pp. 330–338.

[133] M Morsy, S Gouthu, S Orchard, D Thorneycroft, et al. Charting plant interactomes: possibilities and challenges. *Trends Plant Sci* 13 (2008), pp. 183–191.

[134] A R Joyce and B Ø Palsson. The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Biol* 7 (2006), pp. 198–210.

[135] E Marcotte and S Date. Exploiting big biology: integrating large-scale biological data for function inference. *Brief Bioinform* 2 (2001), pp. 363–374.

[136] M A Gilchrist, L A Salter, and A Wagner. A statistical framework for combining and interpreting proteomic datasets. *Bioinformatics* 20 (2004), pp. 689–700.

[137] D Scholtens, M Vidal, and R Gentleman. Local modeling of global interactome networks. *Bioinformatics* 21 (2005), pp. 3548–3557.

[138] H Ge, A J M Walhout, and M Vidal. Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet* 19 (2003), pp. 551–560.

[139] E V Koonin. Systemic determinants of gene evolution and function. *Mol Syst Biol* 1 (2005), p. 2005.0021.

[140] A Kumar, S Agarwal, J A Heyman, S Matson, et al. Subcellular localization of the yeast proteome. *Genes Dev* 16 (2002), pp. 707–719.

[141] W K Huh, J V Falvo, L C Gerke, A S Carroll, et al. Global analysis of protein localization in budding yeast. *Nature* 425 (2003), pp. 686–691.

[142] G R Mishra, M Suresh, K Kumaran, N Kannabiran, et al. Human protein reference database–2006 update. *Nucleic Acids Res* 34 (Database issue) (2006), pp. 411–414.

[143] N Daraselia, A Yuryev, S Egorov, S Novichkova, et al. Extracting human protein interactions from MEDLINE using a full-sentence parser. *Bioinformatics* 20 (2004), pp. 604–611.

[144] H M Müller, E E Kenny, and P W Sternberg. Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS Biol* 2 (2004), e309.

[145] A K Ramani, R C Bunescu, R J Mooney, and E M Marcotte. Consolidating the set of known human protein-protein interactions in preparation for large-scale mapping of the human interactome. *Genome Biol* 6 (2005), R40.

[146] Y Chen and D Xu. Global protein function annotation through mining genome-scale data in yeast *Saccharomyces cerevisiae*. *Nucleic Acids Res* 32 (2004), pp. 6414–6424.

[147] L Kiemer, S Costa, M Ueffing, and G Cesareni. WI-PHI: a weighted yeast interactome enriched for direct physical interactions. *Proteomics* 7 (2007), pp. 932–943.

[148] M G Kann. Protein interactions and disease: computational approaches to uncover the etiology of diseases. *Brief Bioinform* 8 (2007), pp. 333–346.

[149] J Geisler-Lee, N O'Toole, R Ammar, N J Provart, et al. A predicted interactome for *Arabidopsis*. *Plant Physiol* (2007), pp. 317–329.

[150] Y Liu, I Kim, and H Zhao. Protein interaction predictions from diverse sources. *Drug Discov Today* 13 (2008), pp. 409–416.

[151] F Ramírez, A Schlicker, Y Assenov, T Lengauer, et al. Computational analysis of human protein interaction networks. *Proteomics* 7 (2007), pp. 2541–2552.

[152] I Yanai and C DeLisi. The society of genes: networks of functional links between genes from comparative genomics. *Genome Biol* 3 (2002), research0064.

[153] E M Marcotte, M Pellegrini, M J Thompson, T O Yeates, et al. A combined algorithm for genome-wide prediction of protein function. *Nature* 402 (1999), pp. 83–86.

[154] E Sprinzak, S Sattath, and H Margalit. How reliable are experimental protein-protein interaction data? *J Mol Biol* 327 (2003), pp. 919–923.

[155] J C Mellor, I Yanai, K H Clodfelter, J Mintseris, et al. Predictome: a database of putative functional links between proteins. *Nucleic Acids Res* 30 (2002), pp. 306–309.

[156] A Patil and H Nakamura. Filtering high-throughput protein-protein interaction data using a combination of genomic features. *BMC Bioinformatics* 6 (2005), p. 100.

[157] M Gerstein, N Lan, and R Jansen. Integrating interactomes. *Science* 295 (2002), pp. 284–287.

[158] R Jansen, N Lan, J Qian, and M Gerstein. Integration of genomic datasets to predict protein complexes in yeast. *J Struct Funct Genomics* 2 (2002), pp. 71–81.

[159] B Titz, M Schlesner, and P Uetz. What do we learn from high-throughput protein interaction data? *Expert Rev Proteomics* 1 (2004), pp. 111–121.

[160] C von Mering, R Krause, B Snel, M Cornell, et al. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* 417 (2002), pp. 399–403.

[161] J D Han, N Bertin, T Hao, D S Goldberg, et al. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430 (2004), pp. 88–93.

[162] M R Said, T J Begley, A V Oppenheim, D A Lauffenburger, et al. Global network analysis of phenotypic effects: protein networks and toxicity modulation in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 101 (2004), pp. 18006–18011.

[163] A G Fraser and E M Marcotte. A probabilistic view of gene function. *Nat Genet* 36 (2004), pp. 559–564.

[164] S Suthram, T Shlomi, E Ruppin, R Sharan, et al. A direct comparison of protein interaction confidence assignment schemes. *BMC Bioinformatics* 7 (2006), p. 360.

[165] C L Myers and O G Troyanskaya. Context-sensitive data integration and prediction of biological networks. *Bioinformatics* 23 (2007), pp. 2322–2330.

[166] F Browne, H Wang, H Zheng, and F Azuaje. GRIP: A web-based system for constructing Gold Standard datasets for protein-protein interaction prediction. *Source Code Biol Med* 4 (2009), p. 2.

[167] R Jansen and M Gerstein. Analyzing protein function on a genomic scale: the importance of gold-standard positives and negatives for network prediction. *Curr Opin Microbiol* 7 (2004), pp. 535–545.

[168] I Lee, B Lehner, C Crombie, W Wong, et al. A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans*. *Nat Genet* 40 (2008), pp. 181–188.

[169] S Brohée, K Faust, G Lima-Mendez, O Sand, et al. NeAT: a toolbox for the analysis of biological networks, clusters, classes and pathways. *Nucleic Acids Res* 36 (Web Server issue) (2008), W444–451.

[170] C Brun, C Herrmann, and A Guénoche. Clustering proteins from interaction networks for the prediction of cellular functions. *BMC Bioinformatics* 5 (2004), p. 95.

[171] A J Enright, S Van Dongen, and C A Ouzounis. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30 (2002), pp. 1575–1584.

[172] Y Tao, L Sam, J Li, C Friedman, et al. Information theory applied to the sparse Gene Ontology annotation network to predict novel gene function. *Bioinformatics* 23 (2007), pp. 529–538.

[173] B A Shoemaker and A R Panchenko. Deciphering protein-protein interactions. Part II. Computational methods to predict protein and domain interaction partners. *PLoS Comput Biol* 3 (2007), e43.

[174] J Yu and F Fotouhi. Computational approaches for predicting protein-protein interactions: a survey. *J Med Syst* 30 (2006), pp. 39–44.

[175] Y Yamanishi, J P Vert, and M Kanehisa. Protein network inference from multiple genomic data: a supervised approach. *Bioinformatics* 20 Suppl 1 (2004), pp. 363–370.

[176] Y Qi, Z Bar-Joseph, and J Klein-Seetharaman. Evaluation of different biological data and computational classification methods for use in protein interaction prediction. *Proteins* 63 (2006), pp. 490–500.

[177] M S Scott and G J Barton. Probabilistic prediction and ranking of human protein-protein interactions. *BMC Bioinformatics* 8 (2007), p. 239.

[178] B García-Jiménez, D Juan, I Ezkurdia, E Andrés-León, et al. Inference of functional relations in predicted protein networks with a machine learning approach. *PLoS One* 5 (2010), e9969.

[179] J R Bock and D A Gough. Whole-proteome interaction mining. *Bioinformatics* 19 (2003), pp. 125–134.

[180] A Ben-Hur and W S Noble. Kernel methods for predicting protein-protein interactions. *Bioinformatics* 21 Suppl 1 (2005), pp. 38–46.

[181] T-P Nguyen and T-B Ho. An integrative domain-based approach to predicting protein-protein interactions. *J Bioinform Comput Biol* 6 (2008), pp. 1115–1132.

[182] T P Mohamed, J G Carbonell, and M K Ganapathiraju. Active learning for human protein-protein interaction prediction. *BMC Bioinformatics* 11 (2010), S57.

[183] T Aho, H Almusa, J Matilainen, A Larjo, et al. Reconstruction and validation of RefRec: a global model for the yeast molecular interaction network. *PLoS One* 5 (2010), e10662.

[184] S L Lo, C Z Cai, Y Z Chen, and M C Chung. Effect of training datasets on support vector machine prediction of protein-protein interactions. *Proteomics* 5 (2005), pp. 876–884.

[185] J Shen, J Zhang, X Luo, W Zhu, et al. Predicting protein-protein interactions based only on sequences information. *Proc Natl Acad Sci U S A* 104 (2007), pp. 4337–4341.

[186] B-Q Li, K-Y Feng, L Chen, T Huang, et al. Prediction of protein-protein interaction sites by random forest algorithm with mRMR and IFS. *PloS One* 7 (2012), e43927.

[187] E Rodgers-Melnick, M Culp, and S P DiFazio. Predicting whole genome protein interaction networks from primary sequence data in model and non-model organisms using ENTS. *BMC Genomics* 14 (2013), p. 608.

[188] R Pesch and R Zimmer. Complementing the eukaryotic protein interactome. *PloS One* 8 (2013), e66635.

[189] L J Lu, Y Xia, A Paccanaro, H Yu, et al. Assessing the limits of genomic data integration for predicting protein networks. *Genome Res* 15 (2005), pp. 945–953.

[190] Z Wang. Big data mining powers fungal research: recent advances in fission yeast systems biology approaches. *Curr Genet* 63 (2017), pp. 427–433.

[191] T V Vo, J Das, M J Meyer, N A Cordero, et al. A proteome-wide fission yeast interactome reveals network evolution principles from yeasts to human. *Cell* 164 (2016), pp. 310–323.

[192] C J Ryan, A Roguev, K Patrick, J Xu, et al. Hierarchical modularity and the evolution of genetic interactomes across species. *Mol Cell* 46 (2012), pp. 691–704.

[193] J Das, T V Vo, X Wei, J C Mellor, et al. Cross-species protein interactome mapping reveals species-specific wiring of stress response pathways. *Sci Signal* 6 (2013), ra38.

[194] A Roguev, S Bandyopadhyay, M Zofall, K Zhang, et al. Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* 322 (2008), pp. 405–410.

[195] P Beltrao, J C Trinidad, D Fiedler, A Roguev, et al. Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol* 7 (2009), e1000134.

[196] C J Ryan, N J Krogan, P Cunningham, and G Cagney. All or nothing: protein complexes flip essentiality between distantly related eukaryotes. *Genome Biol Evol* 5 (2013), pp. 1049–1059.

[197] E N Koch, M Costanzo, J Bellay, R Deshpande, et al. Conserved rules govern genetic interaction degree across species. *Genome Biol* (2012), R57.

[198] C Ogris, D Guala, and E L L Sonnhammer. FunCoup 4: new species, data, and visualization. *Nucleic Acids Res* 46 (2018), pp. D601–D607.

[199] V Pancaldi, O S Saraç, C Rallis, J R McLean, et al. Predicting the fission yeast protein interaction network. *G3 (Bethesda)* 2 (2012), pp. 453–467.

[200] F Schoeters and P Van Dijck. Protein-protein interactions in *Candida albicans*. *Front Microbiol* 10 (2019), p. 1792.

[201] N A R Gow and B Yadav. Microbe Profile: *Candida albicans*: a shape-changing, opportunistic pathogenic fungus of humans. *Microbiology (Reading)* 163 (2017), pp. 1145–1147.

[202] B Stynen, P Van Dijck, and H Tournu. A CUG codon adapted two-hybrid system for the pathogenic fungus *Candida albicans*. *Nucleic Acids Res* 38 (2010), e184.

[203] F Schoeters, C A Munro, C d'Enfert, and P Van Dijck. A high-throughput *Candida albicans* two-hybrid system. *mSphere* 3 (2018), e00391–18.

[204] T R O'Meara, M J O'Meara, E J Polvi, M R Pourhaghighi, et al. Global proteomic analyses define an environmentally contingent Hsp90 interactome and reveal chaperone-dependent regulation of stress granule proteins and the R2TP complex in a fungal pathogen. *PLoS Biol* 17 (2019), e3000358.

[205] I N Kaneva, I M Sudbery, M J Dickman, and P E Sudbery. Proteins that physically interact with the phosphatase Cdc14 in *Candida albicans* have diverse roles in the cell cycle. *Sci Rep* 9 (2019), p. 6258.

[206] M Urban, A Lysenko, E I Janowska-Sedja, C Rawlings, et al. PHI-Nets: A network resource for Ascomycete fungal pathogens to annotate and identify putative virulence interacting proteins and siRNAs. *Frontiers Microbiol* 10 (2019), p. 2721.

[207] J-P Latgé. *Aspergillus fumigatus* and aspergillosis. *Clin Microbiol Rev* 12 (1999), pp. 310–350.

[208] D W Parry, P Jenkinson, and L McLeod. *Fusarium* ear blight (scab) in small grain cereals—a review. *Plant Pathol* 44 (1995), pp. 207–238.

[209] X Liu, W-H Tang, X-M Zhao, and L Chen. A network approach to predict pathogenic genes for *Fusarium graminearum*. *PloS One* 5 (2010), e13021.

[210] Z-Y Kuo, Y-J Chuang, C-C Chao, F-C Liu, et al. Identification of infection- and defense-related genes via a dynamic host-pathogen interaction network using a *Candida albicans*-zebrafish infection model. *J Innate Immun* 5 (2013), pp. 137–152.

[211] Y-C Wang, C Lin, M-T Chuang, W-P Hsieh, et al. Interspecies protein-protein interaction network construction for characterization of host-pathogen interactions: a *Candida albicans*-zebrafish interaction study. *BMC Syst Biol* 7 (2013), p. 79.

[212] Y-C Wang, I-C Tsai, C Lin, W-P Hsieh, et al. Essential functional modules for pathogenic and defensive mechanisms in *Candida albicans* infections. *Biomed Res Int* 2014 (2014), p. 136130.

[213] C W Remmele, C H Luther, J Balkenhol, T Dandekar, et al. Integrated inference and evaluation of host-fungi interaction networks. *Front Microbiol* 6 (2015), p. 764.

[214] S Ma, Q Song, H Tao, A Harrison, et al. Prediction of protein-protein interactions between fungus (*Magnaporthe grisea*) and rice (*Oryza sativa* L.). *Briefings Bioinform* 20 (2019), pp. 448–456.

[215] S Yang, H Li, H He, Y Zhou, et al. Critical assessment and performance improvement of plant-pathogen protein-protein interaction prediction methods. *Brief Bioinform* 20 (2019), pp. 274–287.

[216] X Xu, T Liu, J Yang, L Chen, et al. The first succinylome profile of *Trichophyton rubrum* reveals lysine succinylation on proteins involved in various key cellular processes. *BMC Genomics* 18 (2017), p. 577.

[217] S Li, B Musungu, D Lightfoot, and P Ji. The interactomic analysis reveals pathogenic protein networks in *Phomopsis longicolla* underlying seed decay of soybean. *Front Genet* 9 (2018), p. 104.

[218] T-Y Wang, F He, Q-W Hu, and Z Zhang. A predicted protein-protein interaction network of the filamentous fungus *Neurospora crassa*. *Mol Biosyst* 7 (2011), pp. 2278–2285.

[219] C Chang, J L Bowman, and E M Meyerowitz. Field guide to plant model systems. *Cell* 167 (2016), pp. 325–339.

[220] J Van Leene, J Hollunder, D Eeckhout, G Persiau, et al. Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*. *Mol Syst Biol* 6 (2010), p. 397.

[221] K Klopffleisch, N Phan, K Augustin, R S Bayne, et al. *Arabidopsis* G-protein interactome reveals connections to cell wall carbohydrates and morphogenesis. *Mol Syst Biol* 7 (2011), p. 532.

[222] Arabidopsis Interactome Mapping Consortium. Evidence for network evolution in an *Arabidopsis* interactome map. *Science* 333 (2011), pp. 601–607.

[223] S A Trigg, R M Garza, A MacWilliams, J R Nery, et al. CrY2H-seq: a massively multiplexed assay for deep-coverage interactome mapping. *Nat Methods* 14 (2017), pp. 819–825.

[224] M Altmann, S Altmann, P A Rodriguez, B Weller, et al. Extensive signal integration by the phytohormone protein network. *Nature* 583 (2020), pp. 271–276.

[225] S C Popescu, G V Popescu, S Bachan, Z Zhang, et al. Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays. *Proc Natl Acad Sci U S A* 104 (2007), pp. 4730–4735.

[226] S C Popescu, G V Popescu, S Bachan, Z Zhang, et al. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes Dev* 23 (2009), pp. 80–92.

[227] A M Jones, Y Xuan, M Xu, R-S Wang, et al. Border control–a membrane-linked interactome of *Arabidopsis*. *Science* 344 (2014), pp. 711–716.

[228] F Zhang, S Liu, L Li, K Zuo, et al. Genome-wide inference of protein-protein interaction networks identifies crosstalk in abscisic acid signaling. *Plant Physiol* 171 (2016), pp. 1511–1522.

[229] J Zhao, Y Lei, J Hong, C Zheng, et al. AraPPINet: an updated interactome for the analysis of hormone signaling crosstalk in *Arabidopsis thaliana*. *Front Plant Sci* 10 (2019), p. 870.

[230] Q-B Yu, G Li, G Wang, J-C Sun, et al. Construction of a chloroplast protein interaction network and functional mining of photosynthetic proteins in *Arabidopsis thaliana*. *Cell Res* 18 (2008), pp. 1007–1019.

[231] Z Ding and D Kihara. Computational identification of protein-protein interactions in model plant proteomes. *Sci Rep* 9 (2019), p. 8740.

[232] M Lin, X Shen, and X Chen. PAIR: the predicted *Arabidopsis* interactome resource. *Nucleic Acids Res* 39 (2011), pp. D1134–D1140.

[233] H Li, S Yang, C Wang, Y Zhou, et al. AraPPISite: a database of fine-grained protein-protein interaction site annotations for *Arabidopsis thaliana*. *Plant Mol Biol* 92 (2016), pp. 105–116.

[234] J Van Leene, J Boruc, G De Jaeger, E Russinova, et al. A kaleidoscopic view of the *Arabidopsis* core cell cycle interactome. *Trends Plant Sci* 16 (2011), pp. 141–150.

[235] Q Lv, Y Lan, Y Shi, H Wang, et al. AtPID: a genome-scale resource for genotype-phenotype associations in *Arabidopsis*. *Nucleic Acids Res* 45 (2017), pp. D1060–D1063.

[236] S De Bodt, S Proost, K Vandepoele, P Rouzé, et al. Predicting protein-protein interactions in *Arabidopsis thaliana* through integration of orthology, gene ontology and co-expression. *BMC Genomics* 10 (2009), p. 288.

[237] I Lee, B Ambaru, P Thakkar, E M Marcotte, et al. Rational association of genes with traits using a genome-scale gene network for *Arabidopsis thaliana*. *Nat Biotechnol* 28 (2010), pp. 149–156.

[238] X Yang, S Yang, H Qi, T Wang, et al. PlaPPISite: a comprehensive resource for plant protein-protein interaction sites. *BMC Plant Biol* 20 (2020), p. 61.

[239] X-G Zhu, J P Lynch, D S LeBauer, A J Millar, et al. Plants *in silico*: why, why now and what?–an integrative platform for plant systems biology research. *Plant Cell Environ* 39 (2016), pp. 1049–1057.

[240] S Struk, A Jacobs, E Sánchez Martín-Fontecha, K Gevaert, et al. Exploring the protein-protein interaction landscape in plants. *Plant Cell Environ* 42 (2019), pp. 387–409.

[241] J F Uhrig. Protein interaction networks in plants. *Planta* 224 (2006), pp. 771–781.

[242] P Braun, S Aubourg, J Van Leene, G De Jaeger, et al. Plant protein interactomes. *Annu Rev Plant Biol* 64 (2013), pp. 161–187.

[243] D Di Silvestre, A Bergamaschi, E Bellini, and P Mauri. Large scale proteomic data and network-based systems biology approaches to explore the plant world. *Proteomes* 6 (2018), p. 27.

[244] S S Sahu, T Weirick, and R Kaundal. Predicting genome-scale *Arabidopsis-Pseudomonas syringae* interactome using domain and interolog-based approaches. *BMC Bioinformatics* 15 Suppl 11 (2014), S13.

[245] A K Biswal, E W McConnell, E G Werth, S-F Lo, et al. The nucleotide-dependent interactome of rice heterotrimeric G-protein α-subunit. *Proteomics* 19 (2019), e1800385.

[246] X Ding, T Richter, M Chen, H Fujii, et al. A rice kinase-protein interaction map. *Plant Physiol* 149 (2009), pp. 1478–1492.

[247] Y-S Seo, M Chern, L E Bartley, M Han, et al. Towards establishment of a rice stress response interactome. *PLoS Genet* 7 (2011), e1002020.

[248] C-L Ho, Y Wu, H-B Shen, N J Provart, et al. A predicted protein interactome for rice. *Rice (N Y)* 5 (2012), p. 15.

[249] S Liu, Y Liu, J Zhao, S Cai, et al. A computational interactome for prioritizing genes associated with complex agronomic traits in rice (*Oryza sativa*). *Plant J* 90 (2017), pp. 177–188.

[250] H Gu, P Zhu, Y Jiao, Y Meng, et al. PRIN: a predicted rice interactome network. *BMC Bioinformatics* 12 (2011), p. 161.

[251] P Zhu, H Gu, Y Jiao, D Huang, et al. Computational identification of protein-protein interactions in rice based on the predicted rice interactome network. *Genomics Proteomics Bioinformatics* (2011), pp. 128–137.

[252] I Lee, Y-S Seo, D Coltrane, S Hwang, et al. Genetic dissection of the biotic stress response using a genome-scale gene network for rice. *Proc Natl Acad Sci U S A* 108 (2011), pp. 18548–18553.

[253] T Lee, T Oh, S Yang, J Shin, et al. RiceNet v2: an improved network prioritization server for rice genes. *Nucleic Acids Res* 43 (2015), W122–W127.

[254] S Lee, T Lee, S Yang, and I Lee. BarleyNet: A network-based functional omics analysis server for cultivated barley, *Hordeum vulgare* L. *Front Plant Sci* 11 (2020), p. 98.

[255] B Musungu, D Bhatnagar, R L Brown, A M Fakhoury, et al. A predicted protein interactome identifies conserved global networks and disease resistance subnetworks in maize. *Front Genet* 6 (2015), p. 201.

[256] J Bhardwaj, I Gangwar, G Panzade, R Shankar, et al. Global *de novo* protein-protein interactome elucidates interactions of drought-responsive proteins in horse gram (*Macrotyloma uniflorum*). *J Proteome Res* 15 (2016), pp. 1794–1809.

[257] J Yue, W Xu, R Ban, S Huang, et al. PTIR: predicted tomato interactome resource. *Sci Rep* 6 (2016), p. 25047.

[258] J Yang, K Osman, M Iqbal, D J Stekel, et al. Inferring the *Brassica rapa* interactome using protein-protein interaction data from *Arabidopsis thaliana*. *Front Plant Sci* 3 (2012), p. 297.

[259] E Fitzek and M Geisler. A predicted interactome for coffee (*Coffea canephora* var robusta). *J Plant Mol Biol Biotechnol* 2 (2011), pp. 34–46.

[260] V Singh Gand Singh and V Singh. Construction and analysis of an interologous protein-protein interaction network of *Camellia sinensis* leaf (TeaLIPIN) from RNA-Seq data sets. *Plant Cell Rep* 38 (2019), pp. 1249–1262.

[261] R Thanasomboon, S Kalapanulak, S Netrphan, and T Saithong. Prediction of cassava protein interactome based on interolog method. *Sci Rep* 7 (2017), p. 17206.

[262] R Thanasomboon, S Kalapanulak, S Netrphan, and T Saithong. Exploring dynamic protein-protein interactions in cassava through the integrative interactome network. *Sci Rep* 10 (2020), p. 6510.

[263] Y-D Ding, J-W Chang, J Guo, D Chen, et al. Prediction and functional analysis of the sweet orange protein-protein interaction network. *BMC Plant Biol* 14 (2014), p. 213.

[264] V Singh, G Singh, and V Singh. TulsiPIN: an interologous protein interactome of *Ocimum tenuiflorum*. *J Proteome Res* 19 (2020), pp. 884–899.

[265] K Vandereyken, B Van Leene Jand De Coninck, and B P A Cammue. Hub protein controversy: taking a closer look at plant stress response hubs. *Front Plant Sci* 9 (2018), p. 694.

[266] L Zhang, J-Y Liu, H Gu, Y Du, et al. *Bradyrhizobium diazoefficiens* USDA 110-*Glycine max* interactome provides candidate proteins associated with symbiosis. *J Proteome Res* 17 (2018), pp. 3061–3074.

[267] R González-Fernández, E Prats, and J V Jorrín-Novo. Proteomics of plant pathogenic fungi. *J Biomed Biotechnol* 2010 (2010), p. 932527.

[268] A Chaiboonchoe, L Ghamsari, B Dohai, P Ng, et al. Systems level analysis of the *Chlamydomonas reinhardtii* metabolic network reveals variability in evolutionary co-conservation. *Mol Biosyst* 12 (2016), pp. 2394–2407.

[269] Y Wang, S Lin, Q Song, K Li, et al. Genome-wide identification of heat shock proteins (Hsps) and Hsp interactors in rice: Hsp70s as a case study. *BMC Genomics* 15 (2014), p. 344.

[270] A Sapkota, X Liu, X-M Zhao, Y Cao, et al. DIPOS: database of interacting proteins in *Oryza sativa*. *Mol Biosyst* 7 (2011), pp. 2615–2621.

[271] S Schuette, B Piatkowski, A Corley, D Lang, et al. Predicted protein-protein interactions in the moss *Physcomitrella patens*: a new bioinformatic resource. *BMC Bioinformatics* 16 (2015), p. 89.

[272] C Ji, X Cao, C Yao, S Xue, et al. Protein–protein interaction network of the marine microalga *Tetraselmis subcordiformis*: prediction and application for starch metabolism analysis. 41 (2014), pp. 1287–1296.

[273] G Zhu, A Wu, X-J Xu, P-P Xiao, et al. PPIM: A protein-protein interaction database for maize. *Plant Physiol* 170 (2016), pp. 618–626.

[274] T Lee, S Lee, S Yang, and I Lee. MaizeNet: a co-functional network for network-assisted systems genetics in *Zea mays*. *Plant Jy* 99 (2019), pp. 571–582.

[275] V M Howick, A J C Russell, T Andrews, H Heaton, et al. The Malaria Cell Atlas: Single parasite transcriptomes across the complete *Plasmodium* life cycle. *Science* 365 (2019), eaaw2619.

[276] D J LaCount. Interactome mapping in malaria parasites: challenges and opportunities. *Methods Mol Biol* 812 (2012), pp. 121–145.

[277] D J LaCount, M Vignali, R Chettier, A Phansalkar, et al. A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* 438 (2005), pp. 103–107.

[278] T Hase, Y Niimura, and H Tanaka. Difference in gene duplicability may explain the difference in overall structure of protein-protein interaction networks among eukaryotes. *BMC Evol Biol* 10 (2010), p. 358.

[279] S Suthram, T Sittler, and T Ideker. The *Plasmodium* protein network diverges from those of other eukaryotes. *Nature* 438 (2005), pp. 108–112.

[280] M D Dyer, T M Murali, and B W. Sobral. Computational prediction of host-pathogen protein-protein interactions. *Bioinformatics* 23 (2007), pp. i159–i166.

[281] S Wuchty. Computational prediction of host-parasite protein interactions between *P. falciparum* and *H. sapiens*. *PloS One* 6 (2011), e26960.

[282] A Rao, M K Kumar, T Joseph, and G Bulusu. Cerebral malaria: insights from host-parasite protein-protein interactions. *Malar J* 9 (2010), p. 155.

[283] A Suratanee and K Plaimas. Heterogeneous network model to identify potential associations between *Plasmodium vivax* and human proteins. *Int J Mol Sci* 21 (2020), p. 1310.

[284] H W Murray, J D Berman, C R Davies, and N G Saravia. Advances in leishmaniasis. *Lancet* 366 (2005), pp. 1561–1577.

[285] J A Pérez-Molina and I Molina. Chagas disease. *Lancet* 391 (2018), pp. 82–94.

[286] A Ponte-Sucre. An overview of *Trypanosoma brucei* infections: an intense host-parasite interaction. *Front Microbiol* 7 (2016), p. 2126.

[287] C R Dos Santos Vasconcelos, T de Lima Campos, and A M Rezende. Building protein-protein interaction networks for *Leishmania* species through protein structural information. *BMC Bioinformatics* 19 (2018), p. 85.

[288] A M Rezende, E L Folador, D de M Resende, and J C Ruiz. Computational prediction of protein-protein interactions in *Leishmania* predicted proteomes. *PloS One* 7 (2012), e51304.

[289] R Watanabe Costa, M F Batista, I Meneghelli, R O Vidal, et al. Comparative analysis of the secretome and interactome of *Trypanosoma cruzi* and *Trypanosoma rangeli* reveals species specific immune response modulating proteins. *Front Immunol* 11 (2020), p. 1774.

[290] H M Kariithi, S Boeren, E K Murungi, J M Vlak, et al. A proteomics approach reveals molecular manipulators of distinct cellular processes in the salivary glands of *Glossina m. morsitans* in response to *Trypanosoma b. brucei* infections. *Parasit Vectors* 9 (2016), p. 424.

[291] E Muriithi, J Kinyua, and S G Nyanjom. *In silico* prediction of protein–protein interaction between *Glossina m. morsitans* (Westwood, 1851) and *Trypanosoma brucei* (Kinetoplastida: Trypanosomatidae). *Bioscience Horizons* 11 (2018), hzy005.

[292] V H Gazestani, C W Yip, N Nikpour, N Berghuis, et al. TrypsNetDB: An integrated framework for the functional characterization of trypanosomatid proteins. *PLoS Negl Trop Dis* 11 (2017), e0005368.

[293] S V Date and C J Stoeckert. Computational modeling of the *Plasmodium falciparum* interactome reveals protein function on a genome-wide scale. *Genome Res* 16 (2006), pp. 542–549.

[294] Y Cuesta-Astroz, A Santos, G Oliveira, and L J Jensen. Analysis of predicted host-parasite interactomes reveals commonalities and specificities related to parasitic lifestyle and tissues tropism. *Front Immunol* 10 (2019), p. 212.

[295] A F Flórez, D Park, J Bhak, B-C Kim, et al. Protein network prediction and topological analysis in *Leishmania major* as a tool for drug target selection. *BMC Bioinformatics* 11 (2010), p. 484.

[296] S R Pavithra, R Kumar, and U Tatu. Systems analysis of chaperone networks in the malarial parasite *Plasmodium falciparum*. *PLoS Comput Biol* 3 (2007), pp. 1701–1715.

[297] A Mitrofanova, S Kleinberg, J Carlton, S Kasif, et al. Predicting malaria interactome classifications from time-course transcriptomic data along the intraerythrocytic developmental cycle. *Artif Intell Med* 49 (2010), pp. 167–176.

[298] C Hillier, M Pardo, L Yu, E Bushell, et al. Landscape of the *Plasmodium* interactome reveals both conserved and species-specific functionality. *Cell Rep* 28 (2019), 1635–1647.e5.

[299] H Huang, E E Winter, H Wang, K G Weinstock, et al. Evolutionary conservation and selection of human disease gene orthologs in the rat and mouse genomes. *Genome Biol* 5 (2004), R47.

[300] T Ravasi, H Suzuki, C V Cannistraci, S Katayama, et al. An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* 140 (2010), pp. 744–752.

[301] Q Cao, X Wang, M Zhao, R Yang, et al. The central role of EED in the orchestration of polycomb group complexes. *Nature Commun* 5 (2014), p. 3127.

[302] J Li, W Zhang, H Yang, D P Howrigan, et al. Spatiotemporal profile of postsynaptic interactomes integrates components of complex brain disorders. *Nat Neurosci* 20 (2017), pp. 1150–1161.

[303] T Zhang, W Du, A F Wilson, S H Namekawa, et al. Fancd2 *in vivo* interaction network reveals a non-canonical role in mitochondrial function. *Sci Rep* 7 (2017), p. 45626.

[304] S Zhang, N A Williamson, and M A Bogoyevitch. Complementary proteomics strategies capture an ataxin-1 interactome in Neuro-2a cells. *Sci Data* 5 (2018), p. 180262.

[305] F Arbogast, J Arnold, P Hammann, L Kuhn, et al. ATG5 is required for B cell polarization and presentation of particulate antigens. *Autophagy* 15 (2 2019), pp. 280–294.

[306] R Pourhaghighi, P E A Ash, S Phanse, F Goebels, et al. BraInMap elucidates the macromolecular connectivity landscape of mammalian brain. *Cell Syst* 10 (2020), 333–350.e14.

[307] C Wan, B Borgeson, S Phanse, F Tu, et al. Panorama of ancient metazoan macromolecular complexes. *Nature* 525 (2015), pp. 339–344.

[308] M A Skinnider, N E Scott, A Prudova, C H Kerr, et al. An atlas of protein-protein interactions across mouse tissues. *Cell* 184 (2021), 4073–4089.e17.

[309] S Yellaboina, D B. Dudekula, and M S Ko. Prediction of evolutionarily conserved interologs in *Mus musculus*. *BMC Genomics* 9 (2008), p. 465.

[310] C J Shin, M J Davis, and M A Ragan. Towards the mammalian interactome: Inference of a core mammalian interaction set in mouse. *Proteomics* 9 (2009), pp. 5256–5266.

[311] X Li, H Cai, J Xu, S Ying, et al. A mouse protein interactome through combined literature mining with multiple sources of interaction evidence. *Amino Acids* 38 (2010), pp. 1237–1252.

[312] G Alanis-Lobato, J S. Möllmann, M H Schaefer, and M A Andrade-Navarro. MIPPIE: the mouse integrated protein-protein interaction reference. *Database (Oxford)* 2020 (2020), baaa035.

[313] W K Kim, C Krumpelman, and E M Marcotte. Inferring mouse gene functions from genomic-scale data using a combined functional network/classification strategy. *Genome Biology* 9 Suppl 1 (2008), S5.

[314] N Guillemin, M Bonnet, C Jurie, and B Picard. Functional analysis of beef tenderness. *J Proteomics* 75 (2011), pp. 352–365.

[315] E Kamau, S G Nyanjom, M Wamalwa, and J Ng'ang'a. Prediction of protein–protein interactions between *Theileria parva* and *Bos taurus* based on sequence homology. *Bioscience Horizons* 9 (2016), hzw006.

[316] F Wang, B Song, X Zhao, Y Miao, et al. Prediction and analysis of the protein-protein interaction networks for chickens, cattle, dogs, horses and rabbits. *Curr Bioinform* 11 (2016), pp. 131–142.

[317] C von Mering, M Huynen, D Jaeggi, S Schmidt, et al. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res* 31 (2003), pp. 258–261.

[318] E Kim, D Bae, S Yang, G Ko, et al. BiomeNet: a database for construction and analysis of functional interaction networks for any species with a sequenced genome. *Bioinformatics* 36 (2020), pp. 1584–1589.

[319] Y Chen, Y Liu, M Du, W Zhang, et al. Constructing a comprehensive gene co-expression based interactome in *Bos taurus*. *PeerJ* 5 (2017), e4107.

[320] Y-T Tao, X-B Ding, J Jin, H-B Zhang, et al. Predicted rat interactome database and gene set linkage analysis. *Database (Oxford)* 2020 (2020), baaa086.

[321] F Wang, M Liu, B Song, D Li, et al. Prediction and characterization of protein-protein interaction networks in swine. *Proteome Sci* 10 (2012), p. 2.

[322] F Dreher, A Kamburov, and R Herwig. Construction of a pig physical interactome using sequence homology and a comprehensive reference human interactome. *Evol Bioinform Online* 8 (2012), pp. 119–126.

[323] J Sprague, L Bayraktaroglu, D Clements, T Conlin, et al. The Zebrafish Information Network: the zebrafish model organism database. *Nucleic Acids Res* 34 (2006), pp. D581–D585.

[324] L-C Tseng, C Zhang, C-M Cheng, H Xu, et al. New classes of mind bomb-interacting proteins identified from yeast two-hybrid screens. *PLoS One* 9 (2014), e93394.

[325] A Alexeyenko, D M Wassenberg, E K Lobenhofer, J Yen, et al. Dynamic zebrafish interactome reveals transcriptional mechanisms of dioxin toxicity. *PloS One* 5 (2010), e10465.

[326] K Waiho, N Afiqah-Aleng, M T M Iryani, and H Fazhan. Protein–protein interaction network: an emerging tool for understanding fish disease in aquaculture. *Rev Aquac* 13 (), pp. 156–177.

[327] H Guo, J Zhang, Y Wang, C Bu, et al. Comparative proteomic analysis of lysine acetylation in fish CIK cells infected with aquareovirus. *Int J Mol Sci* 18 (2017), p. 2419.

[328] H Li, X Huang, Z Zeng, X-X Peng, et al. Identification of the interactome between fish plasma proteins and *Edwardsiella tarda* reveals tissue-specific strategies against bacterial infection. *Int J Biochem Cell Biol* 78 (2016), pp. 260–267.

[329] Y Liu, H Zhang, Y Liu, H Li, et al. Determination of the heterogeneous interactome between *Edwardsiella tarda* and fish gills. *J Proteomics* 75 (2012), pp. 1119–1128.

[330] H Li, Q-F Zhu, X-X Peng, and B Peng. Interactome of *E. piscicida* and grouper liver proteins reveals strategies of bacterial infection and host immune response. *Sci Rep* 7 (2017), p. 39824.

[331] G Kumar, K Hummel, K Noebauer, T J Welch, et al. Proteome analysis reveals a role of rainbow trout lymphoid organs during *Yersinia ruckeri* infection process. *Sci Rep* 8 (2018), p. 13998.

[332] Z Li, X Liu, J Liu, K Zhang, et al. Transcriptome profiling based on protein–protein interaction networks provides a core set of genes for understanding blood immune response mechanisms against *Edwardsiella tarda* infection in Japanese flounder (*Paralichthys olivaceus*). *Dev Comp Immunol* 78 (2018), pp. 100–113.

[333] M Carrera, B Cañas, and J M Gallardo. The sarcoplasmic fish proteome: pathways, metabolic networks and potential bioactive peptides for nutritional inferences. *J Proteomics* 78 (2013), pp. 211–220.

[334] A F Millan-Cubillo, M Martin-Perez, A Ibarz, J Fernandez-Borras, et al. Proteomic characterization of primary cultured myocytes in a fish model at different myogenesis stages. *Sci Rep* 9 (2019), p. 14126.

[335] C Raposo de Magalhães, D Schrama, A P Farinha, D Revets, et al. Protein changes as robust signatures of fish chronic stress: a proteomics approach to fish welfare research. *BMC Genomics* 21 (2020), p. 309.

[336] Z Xu, L Gan, T Li, C Xu, et al. Transcriptome profiling and molecular pathway analysis of genes in association with salinity adaptation in Nile tilapia *Oreochromis niloticus*. *PloS One* 10 (2015), e0136506.

[337] K A Matthews, T C Kaufman, and W M Gelbart. Research resources for *Drosophila*: the expanding universe. *Nat Rev Genet* 6 (2005), pp. 179–193.

[338] L Giot, J S Bader, C Brouwer, A Chaudhuri, et al. A protein interaction map of *Drosophila melanogaster*. *Science* 302 (2003), pp. 1727–1736.

[339] C A Stanyon, G Liu, B A Mangiola, N Patel, et al. A *Drosophila* protein-interaction map centered on cell-cycle regulators. *Genome Biol* 5 (2004), R96.

[340] E Formstecher, S Aresta, V Collura, A Hamburger, et al. Protein interaction mapping: a *Drosophila* case study. *Genome Res* 15 (2005), pp. 376–384.

[341] L Shokri, S Inukai, A Hafner, K Weinand, et al. A comprehensive *Drosophila melanogaster* transcription factor interactome. *Cell Rep* 27 (2019), 955–970.e7.

[342] J S Rees, N Lowe, I M Armean, J Roote, et al. *In vivo* analysis of proteomes and interactomes using Parallel Affinity Capture (iPAC) coupled to mass spectrometry. *Mol Cell Proteomics* 10 (2011), p. M110.002386.

[343] K G Guruharsha, J-F Rual, B Zhai, J Mintseris, et al. A protein complex network of *Drosophila melanogaster*. *Cell* 147 (2011), pp. 690–703.

[344] A A Friedman, G Tucker, R Singh, D Yan, et al. Proteomic and functional genomic landscape of receptor tyrosine kinase and Ras to extracellular signal-regulated kinase signaling. *Sci Signal* 4 (2011), rs10.

[345] H Fukuyama, Y Verdier, Y Guan, C Makino-Okamura, et al. Landscape of protein-protein interactions in *Drosophila* immune deficiency signaling during bacterial challenge. *Proc Natl Acad Sci U S A* 110 (2013), pp. 10717–10722.

[346] Y Kwon, A Vinayagam, X Sun, N Dephoure, et al. The Hippo signaling pathway interactome. *Science* 342 (2013), pp. 737–740.

[347] D Y Rhee, D-Y Cho, B Zhai, M Slattery, et al. Transcription factor networks in *Drosophila melanogaster*. *Cell Rep* 8 (2014), pp. 2031–2043.

[348] A K Gillingham, R Sinka, I L Torres, K S Lilley, et al. Toward a comprehensive map of the effectors of Rab GTPases. *Dev Cell* 31 (2014), pp. 358–373.

[349] A Vinayagam, M M Kulkarni, R Sopko, X Sun, et al. An integrative analysis of the InR/PI3K/Akt network identifies the dynamic response to insulin signaling. *Cell Rep* 16 (2016), pp. 3062–3074.

[350] R Stegeman, P J Spreacker, S K Swanson, R Stephenson, et al. The spliceosomal protein SF3B5 is a novel component of *Drosophila* SAGA that functions in gene expression independent of splicing. *J Mol Biol* 428 (2016), pp. 3632–3649.

[351] A Lunardi, G Di Minin, P Provero, M Dal Ferro, et al. A genome-scale protein interaction profile of *Drosophila* p53 uncovers additional nodes of the human p53 network. *Proc Natl Acad Sci U S A* 107 (2010), pp. 6322–6327.

[352] E Özkan, R A Carrillo, C L Eastman, R Weiszmann, et al. An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. *Cell* 154 (2013), pp. 228–239.

[353] A W Reinke, J Baek, O Ashenberg, and A E Keating. Networks of bZIP protein-protein interactions diversified over a billion years of evolution. *Science* 340 (2013), pp. 730–734.

[354] X-B Ding, J Jin, Y-T Tao, W-P Guo, et al. Predicted *Drosophila* Interactome Resource and web tool for functional interpretation of differentially expressed genes. *Database (Oxford)* 2020 (2020), baaa005.

[355] J Yu, S Pacifico, G Liu, and R L Finley. DroID: the *Drosophila* Interactions Database, a comprehensive resource for annotated gene and protein interactions. *BMC Genomics* 9 (2008), p. 461.

[356] T Murali, S Pacifico, J Yu, S Guest, et al. DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for *Drosophila*. *Nucleic Acids Res* 39 (2011), pp. D736–D743.

[357] X Guo, Y Xu, G Bian, A D Pike, et al. Response of the mosquito protein interaction network to dengue infection. *BMC Genomics* 11 (2010), p. 380.

[358] D Mairiang, H Zhang, A Sodja, T Murali, et al. Identification of new protein interactions between dengue fever virus and its hosts, human and mosquito. *PloS One* 8 (2013), e53535.

[359] J M Doolittle and S M Gomez. Mapping protein interactions between Dengue virus and its human and insect hosts. *PLoS Negl Trop Dis* 5 (2011), e954.

[360] S K Gupta, M Srivastava, Ö Osmanoglu, and T Dandekar. Genome-wide inference of the *Camponotus floridanus* protein-protein interaction network using homologous mapping and interacting domain profile pairs. *Sci Rep* 10 (2020), p. 2334.

[361] R Sumathy, A S K Rao, N Chandrakanth, and V K Gopalakrishnan. *In silico* identification of protein-protein interactions in Silkworm, *Bombyx mori*. *Bioinformation* 10 (2014), pp. 56–62.

[362] L ZM Hu, F Goebels, J H Tan, E Wolf, et al. EPIC: software toolkit for elution profile-based inference of protein complexes. *Nat Methods* 16 (2019), pp. 737–742.

[363] S J Boulton. Combined functional genomic maps of the *C. elegans* DNA damage response. *Science* 295 (2002), pp. 127–131.

[364] L Xu, Y Wei, J Reboul, P Vaglio, et al. BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* 425 (2003), pp. 316–321.

[365] X Xin, J-F Rual, T Hirozane-Kishikawa, D E Hill, et al. Shifted transversal design smart-pooling for high coverage interactome mapping. *Genome Res* 19 (2009), pp. 1262–1269.

[366] N Lenfant, J Polanowska, S Bamps, S Omi, et al. A genome-wide study of PDZ-domain interactions in *C. elegans* reveals a high frequency of non-canonical binding. *BMC Genomics* 11 (2010), p. 671.

[367] T Koorman, D Klompstra, M van der Voet, I Lemmens, et al. A combined binary interaction and phenotypic map of *C. elegans* cell polarity proteins. *Nat Cell Biol* 18 (2016), pp. 337–346.

[368] M Boxem, Z Maliga, N Klitgord, N Li, et al. A protein domain-based interactome network for *C. elegans* early embryogenesis. *Cell* 134 (2008), pp. 534–545.

[369] A B Byrne, M T Weirauch, V Wong, M Koeva, et al. A global analysis of genetic interactions in *Caenorhabditis elegans*. *J Biol* 6 (2007), p. 8.

[370] S M O'Rourke, M D Dorfman, J C Carter, and B Bowerman. Dynein modifiers in *C. elegans*: light chains suppress conditional heavy chain mutants. *PLoS Genet* 3 (2007), e128.

[371] Liquan Cai, Dan Wang, Alfred L. Fisher, and Zhou Wang. Identification of a genetic interaction between the tumor suppressor EAF2 and the retinoblastoma protein (Rb) signaling pathway in *C. elegans* and prostate cancer cells. *Biochem Biophys Res Commun* 447 (2014), pp. 292–298.

[372] J Ceron, J-F Rual, A Chandra, D Dupuy, et al. Large-scale RNAi screens identify novel genes that interact with the *C. elegans* retinoblastoma pathway as well as splicing-related components with synMuv B activity. *BMC Dev Biol* 7 (2007), p. 30.

[373] B Lehner, C Crombie, J Tischler, A Fortunato, et al. Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. *Nat Genet* 38 (2006), pp. 896–903.

[374] *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 282 (1998), pp. 2012–2018.

[375] S Remmelzwaal and M Boxem. Protein interactome mapping in *Caenorhabditis elegans*. *Curr Opin Syst Biol* 13 (2019), pp. 1–9.

[376] S Li, C M Armstrong, N Bertin, H Ge, et al. A map of the interactome network of the metazoan *C. elegans*. *Science* 303 (2004), pp. 540–543.

[377] N Simonis, J-F Rual, A-R Carvunis, M Tasan, et al. Empirically controlled mapping of the *Caenorhabditis elegans* protein-protein interactome network. *Nat Methods* 6 (2009), pp. 47–54.

[378] J S Reece-Hoyes, C Pons, A Diallo, A Mori, et al. Extensive rewiring and complex evolutionary dynamics in a *C. elegans* multiparameter transcription factor network. *Molecular cell* 51 (2013), pp. 116–127.

[379] K C Gunsalus, H Ge, A J Schetter, D S Goldberg, et al. Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature* 436 (2005), pp. 861–865.

[380] P-C Chen, L Ruan, J Jin, Y-T Tao, et al. Predicted functional interactome of *Caenorhabditis elegans* and a web tool for the functional interpretation of differentially expressed genes. *Biol Direct* 15 (2020), pp. 1–11.

[381] X-T Huang, Y Zhu, L L Hang Chan, Z Zhao, et al. An integrative *C. elegans* protein–protein interaction network with reliability assessment based on a probabilistic graphical model. *Mol Biosyst* 12 (2016), pp. 85–92.

[382] K C Gunsalus. A *Caenorhabditis elegans* genetic-interaction map wiggles into view. *J Biol* 7 (3 2008), p. 8.

[383] K C Gunsalus and K Rhrissorrakrai. Networks in *Caenorhabditis elegans*. *Curr Opin Genet Dev* 21 (2011), pp. 787–798.

[384] C M Taylor, K Fischer, S Abubucker, Z Wang, et al. Targeting protein-protein interactions for parasite control. *PloS One* 6 (2011), e18381.

[385] M Berriman, B J Haas, P T LoVerde, R A Wilson, et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 460 (2009), pp. 352–358.

[386] F Luo, M Yin, X Mo, C Sun, et al. An improved genome assembly of the fluke *Schistosoma japonicum*. *PLoS Negl Trop Dis* 13 (2019), e0007612.

[387] J White Bear, T Long, D Skinner, and J H McKerrow. Predictions of novel *Schistosoma mansoni* - human protein interactions consistent with experimental data. *Sci Rep* 8 (2018), p. 13092.

[388] S Caby, L Pagliazzo, J Lancelot, J-M Saliou, et al. Analysis of the interactome of *Schistosoma mansoni* histone deacetylase 8. *PLoS Negl Trop Dis* 11 (2017), e0006089.

[389] Y Moné, B Gourbal, D Duval, L Du Pasquier, et al. A large repertoire of parasite epitopes matched by a large repertoire of host immune receptors in an invertebrate host/parasite model. *PLoS Negl Trop Dis* 4 (2010), e813.

[390] S Castillo-Lara and J F Abril. PlanNET: homology-based predicted interactome for multiple planarian transcriptomes. *Bioinformatics* 34 (2018), pp. 1016–1023.

[391] S Castillo-Lara, E Pascual-Carreras, and J F Abril. PlanExp: intuitive integration of complex RNA-seq datasets with planarian omics resources. *Bioinformatics* 36 (2020), pp. 1889–1895.

[392] J H Konieczka, K Drew, A Pine, K Belasco, et al. BioNetBuilder2.0: bringing systems biology to chicken and other model organisms. *BMC Genomics* 10 Suppl 2 (2009), S6.

[393] M Zhao, T Wang, K J Adamson, K B Storey, et al. Multi-tissue transcriptomics for construction of a comprehensive gene resource for the terrestrial snail *Theba pisana*. *Sci Rep* 6 (2016), p. 20685.

[394] T Hao, Z Zeng, B Wang, Y Zhang, et al. The protein-protein interaction network of eyestalk, Y-organ and hepatopancreas in Chinese mitten crab *Eriocheir sinensis*. *BMC Syst Biol* 8 (2014), p. 39.

[395] T Hao, L Zhao, D Wu, B Wang, et al. The protein-protein interaction network of , *Litopenaeus vannamei* haemocytes. *Front Physiol* 10 (2019), p. 156.

[396] L Jiao, T Dai, M Jin, P Sun, et al. Transcriptome analysis of the hepatopancreas in the *Litopenaeus vannamei* responding to the lead stress. *Biol Trace Elem Res* 199 (2020), pp. 1100–1109.

[397] A E Velázquez-Lizárraga, J L Juárez-Morales, I S Racotta, H Villarreal-Colmenares, et al. Transcriptomic analysis of Pacific white shrimp (*Litopenaeus vannamei*, Boone 1931) in response to acute hepatopancreatic necrosis disease caused by *Vibrio parahaemolyticus*. *PloS One* 14 (2019), e0220993.

[398] X Wang and Y Jin. Predicted networks of protein-protein interactions in *Stegodyphus mimosarum* by cross-species comparisons. *BMC Genomics* 18 (2017), p. 716.

[399] B Müller and U Grossniklaus. Model organisms–A historical perspective. *J Proteomics* 73 (2010), pp. 2054–2063.

[400] S Richards. It's more than stamp collecting: how genome sequencing can unify biological research. *Trends Genet* 31 (2015), pp. 411–421.

[401] J J Russell, J A Theriot, P Sood, W F Marshall, et al. Non-model model organisms. *BMC Biol* 15 (2017), p. 55.

[402] R R da Fonseca, A Albrechtsen, G E Themudo, J Ramos-Madrigal, et al. Next-generation biology: Sequencing and data analysis approaches for non-model organisms. *Mar Genomics* 30 (2016), pp. 3–13.

[403] S Diezmann, M Michaut, R S Shapiro, G D Bader, et al. Mapping the Hsp90 genetic interaction network in *Candida albicans* reveals environmental contingency and rewired circuitry. *PLoS Genet* 8 (2012), e1002562.

[404] S Wuchty, J H Adams, and M T Ferdig. A comprehensive *Plasmodium falciparum* protein interaction map reveals a distinct architecture of a core interactome. *Proteomics* 9 (2009), pp. 1841–1849.

[405] K R Brown and I Jurisica. Unequal evolutionary conservation of human protein interactions in interologous networks. *Genome Biol* 8 (2007), R95.

[406] E L Folador, S S Hassan, N Lemke, D Barh, et al. An improved interolog mapping-based computational prediction of protein-protein interactions with increased network coverage. *Integr Biol (Camb)* 6 (2014), pp. 1080–1087.

[407] F Tebbji, Y Chen, J Richard Albert, K T W Gunsalus, et al. A functional portrait of Med7 and the mediator complex in *Candida albicans*. *PLoS Genet* 10 (2014), e1004770.

[408] J Cui, P Li, G Li, F Xu, et al. AtPID: *Arabidopsis thaliana* protein interactome database–an integrative platform for plant systems biology. *Nucleic Acids Res* 36 (2008), pp. D999–1008.

[409] N M Dheilly, C Adema, D A Raftos, B Gourbal, et al. No more non-model species: the promise of next generation sequencing for comparative immunology. *Dev Comp Immunol* 45 (2014), pp. 56–66.

[410] Q E Fletcher and C Selman. Aging in the wild: Insights from free-living and non-model organisms. *Exp Gerontol* 71 (2015), pp. 1–3.

[411] Mark M. Slabodnick, J. Graham Ruby, Sarah B. Reiff, Estienne C. Swart, et al. The macronuclear Genome of *Stentor coeruleus* reveals tiny introns in a giant cell. *Curr Biol* 27 (2017), pp. 569–575.

[412] X Chen, J R Bracht, A D Goldman, E Dolzhenko, et al. The architecture of a scrambled genome reveals massive levels of genomic rearrangement during development. *Cell* 158 (2014), pp. 1187–1198.

[413] R Ekblom and J Galindo. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity (Edinb)* 107 (2011), pp. 1–15.

[414] D Tagu, J K Colbourne, and N Nègre. Genomic data integration for ecological and evolutionary traits in non-model organisms. *BMC Genomics* 15 (2014), p. 490.

[415] N Rhind, Z Chen, M Yassour, D A Thompson, et al. Comparative functional genomics of the fission yeasts. *Science* 332 (2011), pp. 930–936.

[416] A D Foote, Y Liu, G W C Thomas, T Vinař, et al. Convergent evolution of the genomes of marine mammals. *Nat Genet* 47 (2015), pp. 272–275.

[417] S A Smith, N G Wilson, F E Goetz, C Feehery, et al. Resolving the evolutionary relationships of molluscs with phylogenomic tools. *Nature* 480 (2011), pp. 364–367.

[418] A Ungaro, N Pech, J-F Martin, R J S McCairns, et al. Challenges and advances for transcriptome assembly in non-model species. *PloS One* 12 (2017), e0185020.

[419] N J Wickett, S Mirarab, N Nguyen, T Warnow, et al. Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc Natl Acad Sci U S A* 111 (2014), E4859–E4868.

[420] M Carruthers, A A Yurchenko, J J Augley, C E Adams, et al. *De novo* transcriptome assembly, annotation and comparison of four ecological and evolutionary model salmonid fish species. *BMC Genomics* 19 (2018), p. 32.