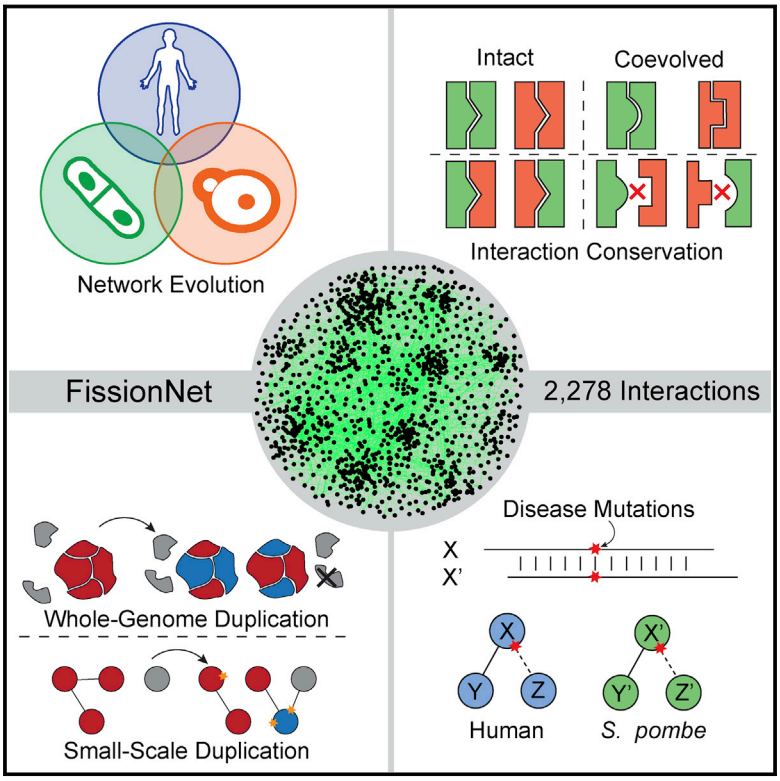
Resource

Cell

A Proteome-wide Fission Yeast Interactome Reveals Network Evolution Principles from Yeasts to Human

Graphical Abstract



Authors

Tommy V. Vo, Jishnu Das, Michael J. Meyer, ..., Jeffrey A. Pleiss, Yu Xia, Haiyuan Yu

Correspondence

haiyuan.yu@cornell.edu

In Brief

FissionNet is a proteome-wide binary interactome network for *S. pombe*. Comparative analyses of FissionNet with protein networks in budding yeast and human reveal how protein networks evolve, principles of gene repurposing following duplication, and insights into species-specific alterations of function.

Highlights

- A proteome-wide network of 2,278 high-quality binary *S. pombe* protein interactions
- Species-specific interaction rewiring among yeasts and human
- Gene duplication modes shape the functional fate of paralogous proteins
- Cross-species interactome mapping uncovers extensive coevolution

Accession Numbers

GSE74354

Vo et al., 2016, Cell *164*, 310–323
January 14, 2016 ©2016 Elsevier Inc.
http://dx.doi.org/10.1016/j.cell.2015.11.037

CellPress

Cell

Resource

A Proteome-wide Fission Yeast Interactome Reveals Network Evolution Principles from Yeasts to Human

Tommy V. Vo,[1,2,3,14] Jishnu Das,[1,2,14] Michael J. Meyer,[1,2,4,14] Nicolas A. Cordero,[2] Nurten Akturk,[2] Xiaomu Wei,[2,5] Benjamin J. Fair,[3] Andrew G. Degatano,[2] Robert Fragoza,[1,2,3] Lisa G. Liu,[2] Akihisa Matsuyama,[6] Michelle Trickey,[7] Sachi Horibata,[8] Andrew Grimson,[3] Hiroyuki Yamano,[7] Minoru Yoshida,[6] Frederick P. Roth,[9,10,11,12] Jeffrey A. Pleiss,[3] Yu Xia,[9,13] and Haiyuan Yu[1,2,*]

[1]Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA
[2]Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY 14853, USA
[3]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA
[4]Tri-Institutional Training Program in Computational Biology and Medicine, New York, NY 10065, USA
[5]Department of Medicine, Weill Cornell College of Medicine, New York, NY 10021, USA
[6]Chemical Genomics Research Group, RIKEN Center for Sustainable Resource Center, Wako, Saitama 351-0198, Japan
[7]University College London Cancer Institute, Paul O'Gorman Building, 72 Huntley Street, London WC1E 6BT, UK
[8]Department of Biomedical Sciences, Baker Institute for Animal Health, Cornell University, Ithaca, NY 14853, USA
[9]Center for Cancer Systems Biology and Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[10]Donnelly Centre and Departments of Molecular Genetics and Computer Science, University of Toronto, Toronto, ON M5S 3E1, Canada
[11]Canadian Institute for Advanced Research, Toronto, ON M5G 1Z8, Canada
[12]Lunenfeld-Tanenbaum Research Institute, Mt. Sinai Hospital, Toronto, ON M5G 1X5, Canada
[13]Department of Bioengineering, Faculty of Engineering, McGill University, Montreal, QC H3A 0C3, Canada
[14]Co-first author
*Correspondence: haiyuan.yu@cornell.edu
http://dx.doi.org/10.1016/j.cell.2015.11.037

SUMMARY

Here, we present FissionNet, a proteome-wide binary protein interactome for *S. pombe*, comprising 2,278 high-quality interactions, of which ~50% were previously not reported in any species. FissionNet unravels previously unreported interactions implicated in processes such as gene silencing and pre-mRNA splicing. We developed a rigorous network comparison framework that accounts for assay sensitivity and specificity, revealing extensive species-specific network rewiring between fission yeast, budding yeast, and human. Surprisingly, although genes are better conserved between the yeasts, *S. pombe* interactions are significantly better conserved in human than in *S. cerevisiae*. Our framework also reveals that different modes of gene duplication influence the extent to which paralogous proteins are functionally repurposed. Finally, cross-species interactome mapping demonstrates that coevolution of interacting proteins is remarkably prevalent, a result with important implications for studying human disease in model organisms. Overall, FissionNet is a valuable resource for understanding protein functions and their evolution.

INTRODUCTION

Proteins function primarily by physically interacting with other proteins. Gain or loss of these interactions within an organism can modulate protein functions and disease states (Sahni et al., 2015; Wei et al., 2014). The importance of protein interactions to our understanding of fundamental biological processes has spurred the mapping of protein interactome networks for several organisms (*Arabidopsis* Interactome Mapping Consortium, 2011; Giot et al., 2003; Rolland et al., 2014; Stelzl et al., 2005; Yu et al., 2008). However, the budding yeast *Saccharomyces cerevisiae* remains the only eukaryotic organism for which a high-coverage binary protein interactome has been mapped by systematic interrogation of pairwise combinations of all proteins in triplicate (Yu et al., 2008). Here, we present FissionNet, a high-coverage proteome-wide protein interactome network generated for the fission yeast *Schizosaccharomyces pombe*.

We compared FissionNet with the only other proteome-scale eukaryotic interactomes available (> 50% of all protein pairs screened), the interactome networks of *S. cerevisiae* and human. Surprisingly, we find that FissionNet is more similar to the human network than it is to that of *S. cerevisiae*. Furthermore, among interactions involving conserved proteins, there is significant species-specific rewiring that is not completely determined by overall sequence similarity of orthologs. Instead, we identify several other determinants of interaction conservation, including local network constraints and conservation of interacting protein domains. Also, by comparing FissionNet with the proteome-wide interactome of *S. cerevisiae*, we are able to ascertain how gene duplication events influence the process by which paralogs acquire novel functions.

S. pombe is an important model organism for studying fundamental biological processes such as RNA splicing, cell-cycle regulation, RNAi, and centromeric maintenance, which are conserved in metazoans but divergent in budding yeast (Wood

et al., 2002). We use FissionNet to unveil previously unreported protein associations between gene regulatory factors involved in pre-mRNA splicing and silencing of stress-response genes and at pericentromeric regions, illustrating the value of our network as a proteome-scale resource to understand biological processes.

RESULTS

A Proteome-wide High-Coverage Binary Protein Interactome Map of *S. pombe*

To generate a proteome-wide interactome network for *S. pombe*, which we call FissionNet, we systematically tested all pairwise combinations of proteins encoded by 4,989 *S. pombe* genes (corresponding to > 99% of all *S. pombe* coding genes) using our high-quality yeast two-hybrid (Y2H) assay, the same pipeline that we used to generate the budding yeast and human interactome networks (Figure S1A) (Yu et al., 2008, 2011). Extensive screenings in triplicate (a total of ~75 million protein pairs) yielded 2,278 interactions between 1,305 proteins, of which 2,130 (93.5%) have not been previously reported in *S. pombe* (Figure 1A) (Das and Yu, 2012). Furthermore, FissionNet contains 1,034 interactions that have not been reported between orthologs in any other species before. Of these, 142 interactions involve *S. pombe* proteins that both have human orthologs, but at least one does not have a *S. cerevisiae* ortholog and, hence, cannot be studied in *S. cerevisiae*. Thus, FissionNet provides a valuable repertoire of biological insights.

To assess the sensitivity and specificity of our Y2H assay (Yu et al., 2008), we constructed a positive reference set (PRS) consisting of 93 well-validated *S. pombe* interactions from the literature and a negative reference set (NRS) of 168 random *S. pombe* protein pairs that are not known to interact in the literature and whose orthologs in other species are also not known to interact (see Supplemental Experimental Procedures). We performed Y2H and protein complementation assay (PCA) (Das et al., 2013; Yu et al., 2008) to test what fraction of the PRS, NRS, and a random sample of 220 FissionNet interactions can be detected using orthogonal methods (Figure 1B). We found that the detection rates of the PRS and FissionNet interactions are indistinguishable from each other and are significantly higher than that of the NRS (Figure 1B; >15% difference in detection rates between the PRS and NRS for both assays, $p < 10^{-3}$, *Z* test). The robust validation rates of FissionNet interactions by an orthogonal assay confirm the high quality of the network. Furthermore, although it has been speculated that Y2H interactions involving proteins with many interaction partners (hubs) could be of low quality (Bader et al., 2004), we found that the validation rate by PCA of hub interactions is the same as the overall PCA validation rate for FissionNet (Figure 1B; $p = 0.34$, *Z* test), confirming that FissionNet interactions involving hubs are of high quality.

Biological relationships between interacting proteins in FissionNet were assessed by measuring similarities in protein localization, functional annotations, and expression profiles (see Supplemental Experimental Procedures). We found that FissionNet interactions are significantly enriched for protein pairs that are co-localized, functionally similar, and encoded by coexpressed genes relative to random expectation (Figures 1C–1E and S1B–S1H; $p < 0.05$ in all three cases using a Kolmogorov-Smirnov (KS) test for coexpression and *Z* test for co-localization and functional similarity). Furthermore, the enrichment of these interactions for all three categories is similar to that of literature-curated binary interactions. These results confirm that FissionNet interactions are functionally relevant in vivo. We illustrate this by focusing on two previously unreported interactions: Tas3-Hhp1 and Atf1-Cid12, and their potential roles in gene silencing.

FissionNet Provides Insights into Functions of Proteins and Interactions

The regulation of centromeric silencing is a well-conserved process in *S. pombe* and metazoans but is divergent from that in *S. cerevisiae* (Holoch and Moazed, 2015). FissionNet revealed a previously unidentified interaction between Tas3 and Hhp1 that we confirmed in vivo (Figures 1F and 1G). Tas3 is a component of the RNA-induced transcriptional silencing (RITS) complex that mediates gene silencing at *S. pombe* centromeres (Verdel et al., 2004). Hhp1 is a conserved mitotic checkpoint kinase (Johnson et al., 2013) not known to be involved in centromeric silencing. In *S. pombe* cells where the *ura4*$^+$ reporter gene was inserted at the centromere inner repeats of chromosome 1 (*imr1R*) (Verdel et al., 2004), we find that *hhp1Δ* confers loss of silencing at the centromere, similar to *tas3Δ* cells (Figure 1H). Furthermore, levels of endogenous centromeric transcripts are elevated in *hhp1Δ* cells (Figure S1I). Moreover, loss of *hhp1* leads to a decrease in the dimethylation of histone 3 lysine 9 (H3K9me2) at the centromere (Figure S1J). These results show that Hhp1 is involved in centromeric silencing.

We also identified a previously unreported interaction between the transcription factor Atf1 and the polyadenylation polymerase Cid12 (Figure 2A). Atf1 mediates transcriptional responses to stresses such as high temperatures (Shiozaki and Russell, 1996). At *S. pombe* centromeres, Cid12 is a core component of the RNA-directed RNA-polymerase complex (RDRC) (Motamedi et al., 2004). The RDRC is responsible for generating double-stranded RNAs, a key step for Dcr1-dependent centromeric silencing. Interestingly, it has been reported that Dcr1 transcriptionally represses the Atf1-target genes *hsp16* and *hsp104* under non-stressed conditions (Woolcock et al., 2012).

Pull-down experiments confirm the interaction of Atf1 and Cid12 in *S. pombe* (Figure 2B), and *cid12Δ* cells grown under non-stressed conditions show elevated mRNA levels of *hsp16* and *hsp104* as compared to wild-type cells, similar to *dcr1Δ* cells (Figure 2C). Additionally, double-mutant *cid12Δ dcr1Δ* cells do not exhibit more drastic transcript accumulation than the single-deletion mutants, suggesting both genes function in the same pathway (Figure 2C). Together, these results suggest that Cid12 may be involved in repressing aberrant gene expression of Atf1-target genes.

Next, we identified two Cid12 mutations, lysine-213 to isoleucine (Cid12^{K213I}) and aspartic acid-260 to valine (Cid12^{D260V}), that disrupt the interaction of Cid12 with Atf1 while preserving interactions within the RDRC complex (Figure 2D; see Supplemental Experimental Procedures). Exogenous expression of wild-type Cid12 in *cid12Δ* cells enables the transcriptional

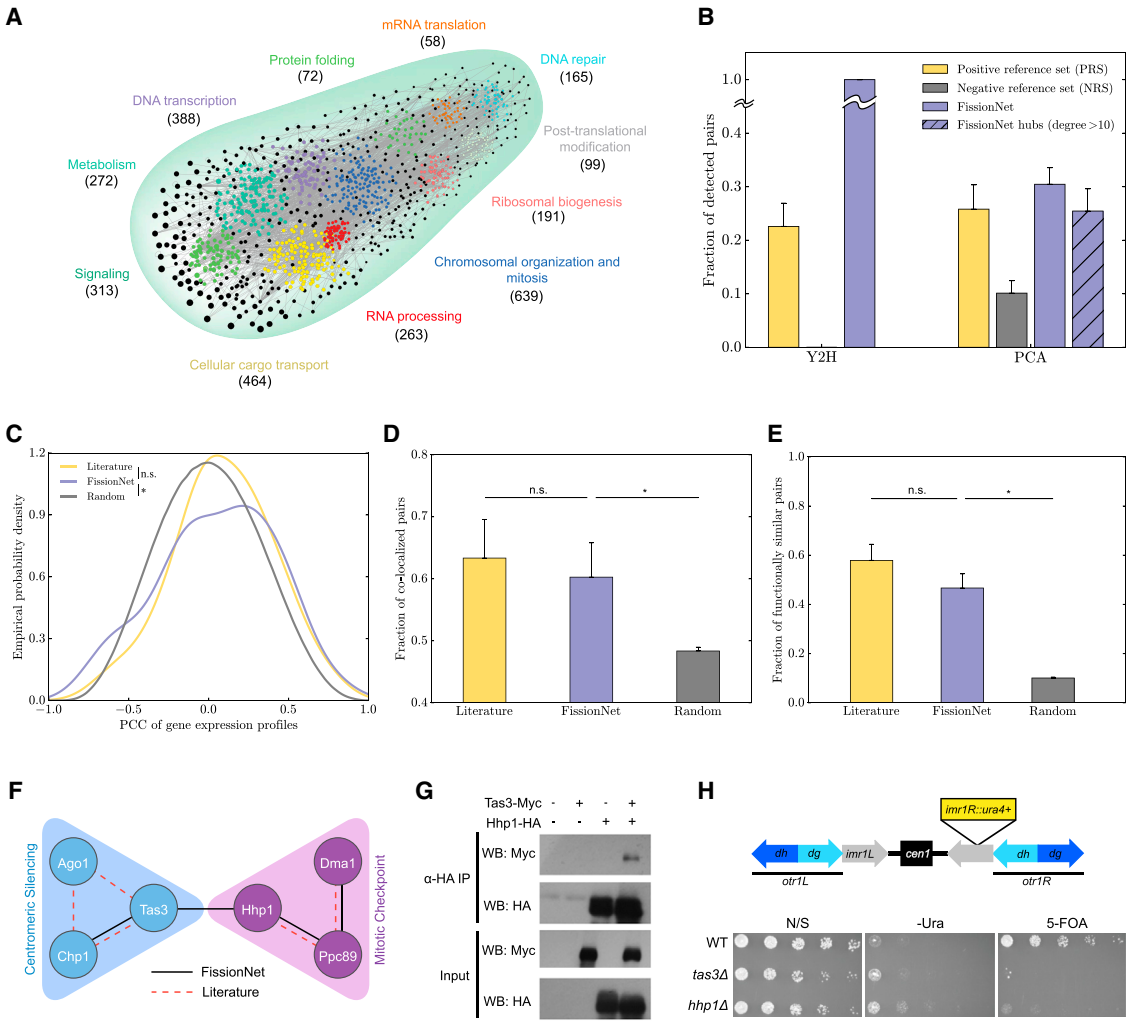


Figure 1. A Proteome-wide Binary Protein Interactome Map of *S. pombe*

(A) Network representation of FissionNet. Proteins are color-grouped based on PomBase GO slim categories. The number of FissionNet interactions per group is indicated.

(B) Y2H and PCA detection rates of the PRS, NRS, FissionNet, and FissionNet hub interactions.

(C) Pearson correlation coefficient (*PCC*) distribution of gene expression profiles of interacting and all random protein pairs.

(D) Enrichment of co-localized protein pairs.

(E) Enrichment of protein pairs sharing similar functions.

(F) Subnetwork of Tas3 and Hhp1 in FissionNet.

(G) Coimmunoprecipitation of Tas3-myc and Hhp1-HA in vivo.

(H) Centromeric silencing assay of *tas3Δ* and *hhp1Δ* cells. A schematic of the *imr1R* region with the *ura4*$^+$ reporter gene is shown. WT denotes wild-type. Data are shown as measurements + SE. * denotes significant ($p < 0.05$); n.s. denotes not significant.

See also Figure S1 and Tables S1 and S2.

repression of *hsp16* and *hsp104*. In stark contrast, neither mutant can repress gene expression (Figure 2E). The mutant phenotype is not due to complete loss of protein caused by destabilization because these Cid12 mutant proteins express in *S. pombe* cells (Figure S2A). Furthermore, in *cid12Δ* cells where the *ura4*$^+$ reporter gene was inserted at the centromeric *imr1R*, we find that exogenous expression of either Cid12 wild-type or mutants equally permit the silencing of the *ura4*$^+$ reporter (Figure 2F). Thus, we show that Cid12 has dual roles in regulating the expression of heat-shock genes and the centromere. Importantly, the roles can be selectively uncoupled via specific disruption of the Atf1-Cid12 interaction. These examples illustrate the usefulness of FissionNet as a resource to uncover areas of biological inquiry.

Comparative Network Analyses Reveal Species-Specific Conservation of Interactions

High-quality protein interactome networks have previously been reported in budding yeast (Yu et al., 2008) and human (Rolland et al., 2014). A fundamental question, which can be addressed

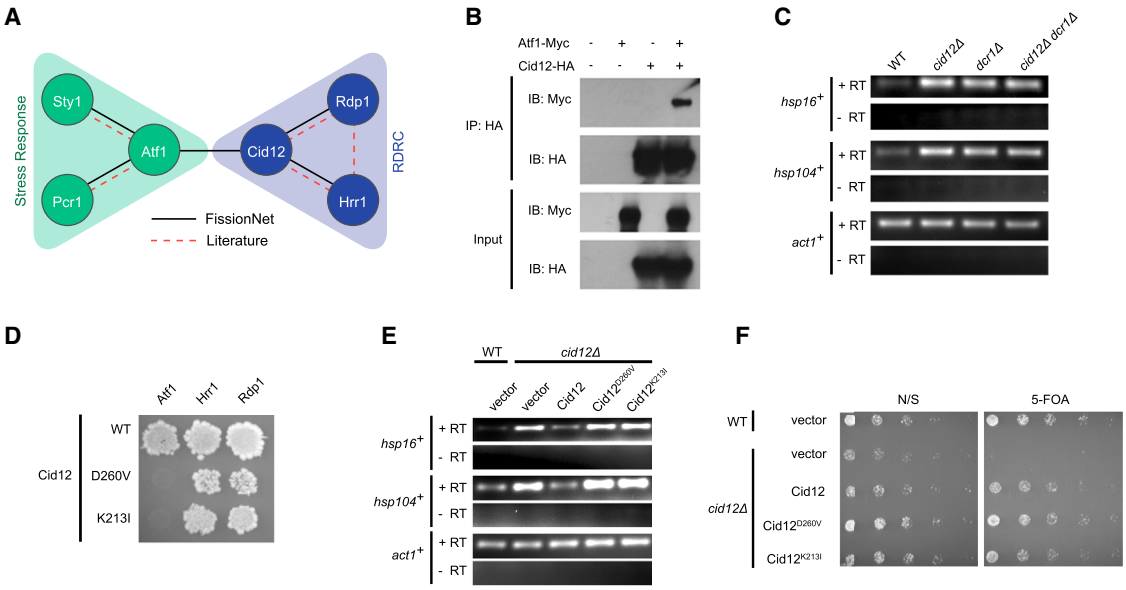


Figure 2. Atf1-Cid12 Interaction Mediates Silencing at Heat-Shock Genes

(A) Subnetwork of Atf1 and Cid12 in FissionNet.

(B) Coimmunoprecipitation of Atf1-myc and Cid12-HA in vivo.

(C) Semiquantitative real-time PCR (semi qRT-PCR) shows *hsp16* and *hsp104* transcript levels in deletion strains.

(D) Y2H confirms Cid12 mutants cannot interact with Atf1 but maintain interactions with Hrr1 and Rdp1.

(E) Semi qRT-PCR shows that the Cid12 mutants in *cid12Δ* cells do not restore the repression of *hsp16* or *hsp104*.

(F) Centromeric silencing assay shows that Cid12 mutants retain centromeric silencing function.

−RT, no reverse transcriptase; +RT, with reverse transcriptase; $Act1^{+}$ serves as loading control; WT, wild-type.

See also Figure S2.

with FissionNet and these networks, is how protein-protein interactions have evolved and whether this trend mirrors gene-level evolution. From sequence-based phylogenetic analyses, the two yeasts are less divergent from each other than either yeast is from human (Figure 3A) (Sipiczki, 2000). Additionally, the two yeasts share a greater fraction of protein-coding genes than either yeast does with human (Figures S3A and S3B).

To calculate interaction conservation, we considered only those interactions that have the potential to be conserved, i.e., the two interacting proteins in the reference species have orthologs in the other species. However, directly calculating the overlap between sets of interactions obtained from the literature would be erroneous because currently available interactomes are incomplete and are derived from assays with varied and often unreported false-positive and false-negative rates (Yu et al., 2008). Therefore, to accurately estimate the underlying interaction conservation fractions, we required interactomes of all species to be derived from the same experimental assay. Since interactomes in budding yeast (Yu et al., 2008) and human (Rolland et al., 2014) have been generated using our version of Y2H (Figures S3C and S3D), we were able to compare FissionNet to these interactome networks to measure the observed extent of interaction conservation. We developed a rigorous Bayesian framework that incorporates both the false-positive and false-negative rates of our Y2H assay to estimate the underlying interaction conservation fraction from the observed fraction for each pair of species (see Supplemental Experimental Procedures). Surprisingly, we find that interaction conservation follows a completely different trend from gene conservation (Figures 3B, S3E, and S3F). While only ~40% of *S. pombe* interactions are conserved in *S. cerevisiae* (of the 1,331 interactions where both proteins have *S. cerevisiae* orthologs and were pairwise retested using our Y2H assay), ~65% of *S. pombe* interactions are conserved in human (of the 652 interactions where both proteins have human orthologs and were pairwise retested using our Y2H assay) (Figure 3B; $p = 1.4 \times 10^{-4}$, *Z* test). However, when using budding yeast as the reference species, the fraction of conserved interactions is as high in fission yeast as in human, comparable to the fraction conserved between fission yeast and human (Figure 3B). We were able to recapitulate these results using interaction datasets generated by other assays (Figures S3G–S3I; >1.5-fold difference between fission yeast interactions conserved in budding yeast and human; $p < 10^{-3}$ in all cases, *Z* test; see Supplemental Experimental Procedures). Thus, our results suggest that a large fraction of interactions are conserved between human and *S. pombe* but have been lost specifically in the *S. cerevisiae* lineage.

One possible explanation for these surprising results is that fission yeast proteins that are conserved in human could have higher overall sequence similarity than those that are conserved in budding yeast. However, we find that proteins in interactions that have the potential to be conserved based on orthology are actually slightly more similar in sequence between the two

yeasts than between *S. pombe* and human (Figure 3C; $p < 10^{-5}$, *U* test; see Supplemental Experimental Procedures).

Another possibility is that the observed difference primarily arises from interactions involving proteins that are conserved between fission yeast and human but lost in budding yeast. To test this, we first focused on proteins that are conserved in all three species. We still find that ~20% more interactions are conserved between *S. pombe* and human as compared to between the two yeasts (Figures 3D, S3J, and S3K; $p < 0.05$, *Z* test).

We next explored the conservation of interactions involved in various biological processes as defined by the Gene Ontology (GO) (Ashburner et al., 2000). We find wide variation in species-specific interaction conservation among different processes (Figures 3E and S3L–S3N). We show that *S. pombe* interactions are more conserved in human than in *S. cerevisiae* for 10 out of 13 GO Slim categories containing $\geq$ 50 interactions (Figure 3E; $p < 0.05$, as marked, *Z* test). The same trend is observed with GO Slim categories containing $\geq$ 30 or $\geq$ 75 interactions (Figures S3L and S3N). Some of these categories, such as "chromosomal organization," "chromosome segregation," and "cell cycle," are far better conserved in human than in *S. cerevisiae*, and accordingly *S. pombe* has been used as a model organism for studying these processes (Wood et al., 2002). Furthermore, considering GO Slim categories that are well conserved in all three species (using cutoffs of $\geq$ 50, 100, and 200 genes annotated per species), we find that the conservation of *S. pombe* interactions in these core biological processes is also higher in human than in *S. cerevisiae* (Figures 3F and S3O; $p < 10^{-3}$, *Z* test). Overall, these results suggest that insights gained from FissionNet may be widely applicable to the study of human biology across many important cellular processes.

We validated three cases of previously unreported functional conservation between fission yeast and human proteins. Uncharacterized *S. pombe* factors Srrm1, SPAC30D11.14C, and SPAC1952.06C interact with known splice factors Srp1, Usp104, and Cwf15, respectively (Figure 3G). Although these proteins have no orthologs in *S. cerevisiae*, they are orthologous to human SRRM1, KIAA0907, and CTNNBL1, respectively. Interestingly, all three human orthologs have been implicated in pre-mRNA splicing or were found to associate with spliceosomal factors in human (Blencowe et al., 1998; Hegele et al., 2012; Rolland et al., 2014). We used DNA microarrays to measure changes in the splicing of every known intron in the *S. pombe* deletion mutants. The loss of *srrm1*, *SPAC30D11.14C*, or *SPAC1952.06C* results in widespread splicing defects, confirming the roles for these proteins in the splicing pathway (Figure 3H). Moreover, Srrm1 and Srp1 share many gene targets, suggesting that the interacting proteins are functionally related (Figure S3P). Notably, an analysis of the introns whose splicing is affected by *srrm1* deletion shows a strong enrichment for introns with weak splice site signals (Figure S3Q). This is consistent with previous findings that human *SRRM1* affects splice site selection by binding to exonic splicing enhancers and facilitating interactions between spliceosomal proteins (Blencowe et al., 1998). These results highlight the utility of FissionNet to reveal proteins that are functionally conserved between *S. pombe* and human.

Determinants of Interaction Conservation

Previous studies have shown that increased protein sequence similarity facilitates conservation of protein interactions (Matthews et al., 2001). Indeed, we also find a positive correlation between sequence similarity of proteins and the fraction of their associated interactions conserved between *S. pombe* and human or *S. cerevisiae*, demonstrating a proteome-scale dependence of protein sequence and function (Figure 4A; $R^2_{S.p\text{-}H.s} = 0.948$ and $R^2_{S.p\text{-}S.c} = 0.976$). However, protein interaction conservation is not completely dependent on overall sequence similarity, as we find many instances of conserved interactions involving proteins with low overall sequence similarity (<40%) with their orthologs (Figure 4A; 40% and 13% of 116 interactions in human and 196 interactions in *S. cerevisiae*, respectively). To investigate whether certain highly conserved domains in these proteins play an important role in interaction conservation, we inferred protein interaction domains from co-crystal structures of 124 human interactions conserved in *S. pombe* and 293 conserved in *S. cerevisiae*. We find that the sequence similarity within protein interaction domains tends to be higher than in other domains for interactions conserved between fission yeast and human (Figure 4B; 7.0% higher, $p = 0.012$, *U* test). For instance, the human DR1-DRAP1 heterodimer is orthologous to the protein pair Ncb2 and Dpb3 in *S. pombe.* While the overall sequence similarity of the orthologs is quite low (0.58 and 0.51, respectively), the interaction is conserved in fission yeast. Moreover, we also find that the proteins can interact with the orthologs of their native interaction partner (Figure 4C). On the basis of a crystal structure of the human DR1-DRAP1 complex, we were able to determine the interaction domains of these proteins (Figure 4D) (Kamada et al., 2001). The sequence similarity within these domains in DR1 and DRAP1 with their fission yeast orthologs is 0.78 and 0.80, respectively, while the conservation outside of these interaction domains is only 0.45 and 0.38. Thus, the basis for this high degree of functional conservation is likely dependent on the interaction domains.

Strikingly, interaction conservation is nearly three times higher between *S. pombe* and human than between the two yeasts at low levels of overall sequence similarity (Figure 4A; at <40% similarity, $p = 0.030$, *Z* test). As sequence similarity approaches 100%, interaction conservation converges. Therefore, for the vast majority of interactions corresponding to proteins with lower sequence similarity to their orthologs, our results strongly suggest that species-specific factors, independent of overall protein sequence similarity, influence conservation of protein-protein interactions.

We then sought to explore other factors that could explain the basis of interaction conservation. First, we used ClusterOne (Nepusz et al., 2012) to detect topological protein clusters in FissionNet (see Supplemental Experimental Procedures). We find that intra-cluster FissionNet interactions are more than three times more likely to be conserved in both budding yeast and human than inter-cluster interactions (Figures 4E and S4A; $p < 0.05$ for both organisms, *Z* test). Next, we examined biological processes defined by GO (Ashburner et al., 2000) and observed the same trend (Figures 4F and S4B; $p < 10^{-3}$ for both organisms, *Z* test). On the basis of genetic interactions, it has been earlier hypothesized that while individual functional modules

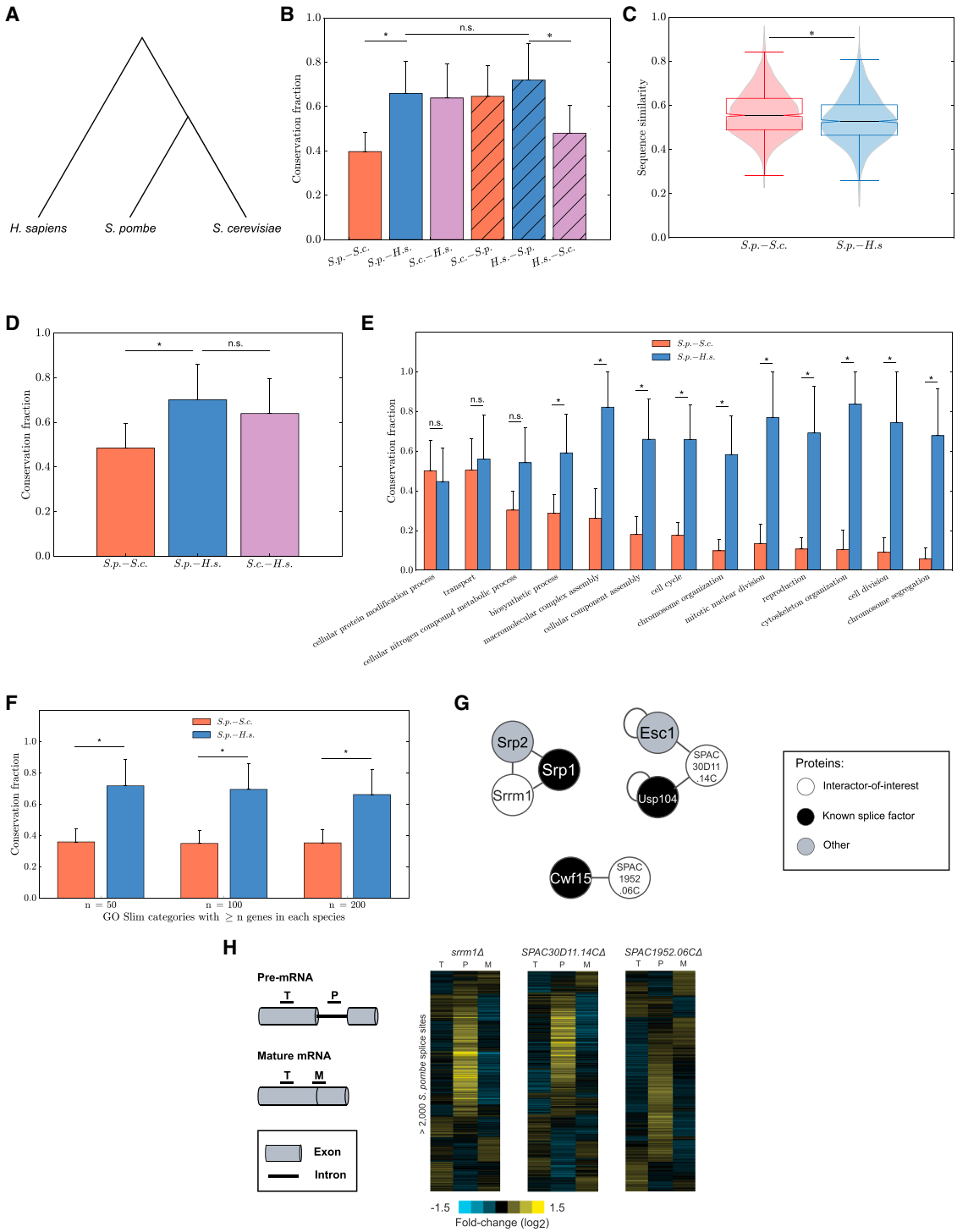


Figure 3. *S. pombe* Protein Interactions Are More Conserved in Human than in *S. cerevisiae*

(A) Sequence-based phylogeny dendrogram of *S. pombe* (*S.p.*), *S. cerevisiae* (*S.c.*), and human (*H.s.*).

(B) Interaction conservation between reference-query species.

(C) Sequence conservation for ortholog pairs that could be conserved between *S.p.*-*S.c.* and *S.p.*-*H.s.*

(D) Interaction conservation between reference-query species for proteins that are conserved in all three species.

(E) Interaction conservation in GO Slim categories with at least 50 interactions.

(F) Interaction conservation among GO Slim categories that are conserved in all three species.

(legend continued on next page)

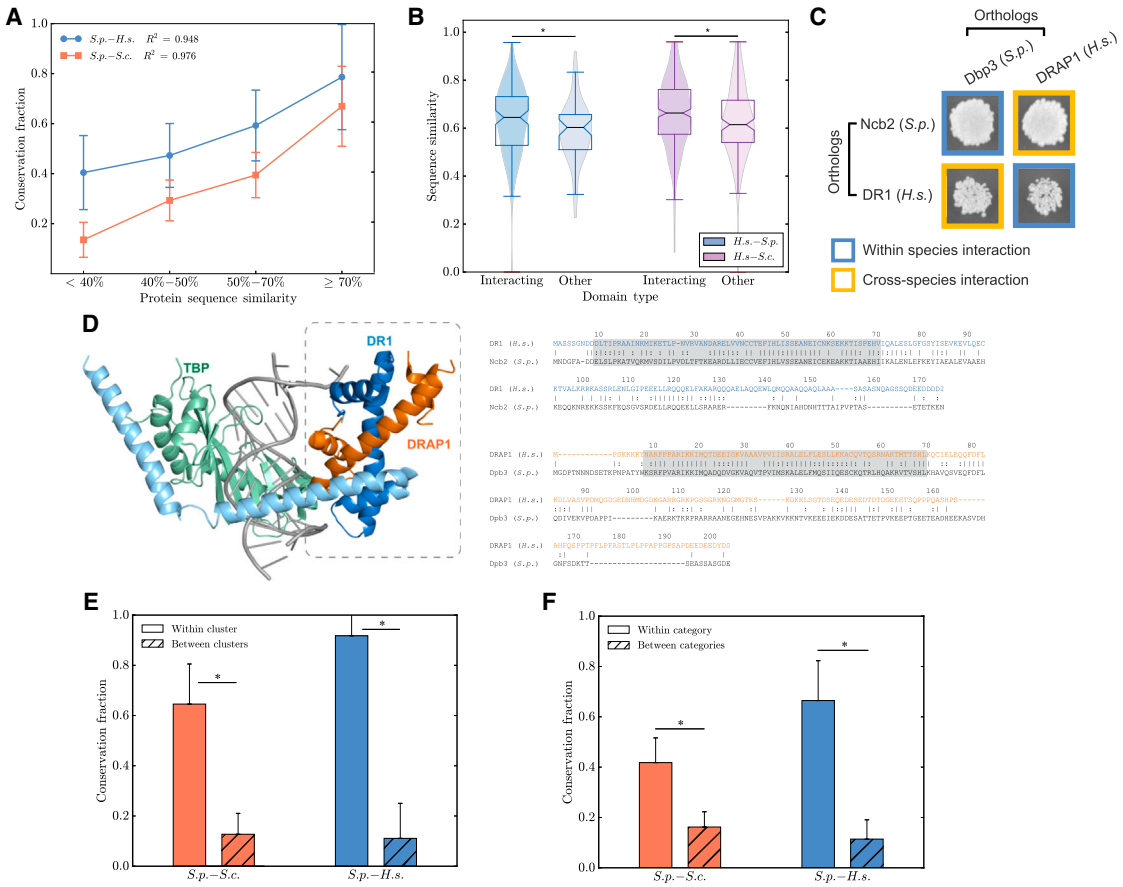


Figure 4. Determinants of Interaction Conservation

(A) Interaction conservation as a function of overall protein sequence similarity.

(B) Sequence similarity within protein interaction domains and other domains for interactions conserved between yeasts and human.

(C) Y2H confirms the interactions of human (*H.s.*) DRAP1-DR1, the orthologous *S. pombe* (*S.p.*) Dpb3-Ncb2, and the cross-species interactions.

(D) Crystal structure of human DR1-DRAP1. Boxed region highlights interaction domains. Gray shaded regions denote aligned interaction domain sequences.

(E) Interaction conservation within and across topological clusters.

(F) Interaction conservation within and across GO categories. Data are shown as measurements + SE. * denotes significant ($p < 0.05$); n.s. denotes not significant. Abbreviations are *S. pombe* (*S.p.*), *S. cerevisiae* (*S.c.*), and human (*H.s.*).

See also Figure S4.

are conserved, inter-modular connectivity could be rewired across evolution (Roguev et al., 2008). In this study, we provide direct molecular-level evidence on a proteome scale that, although interactions within modules tend to be conserved across evolution, the cross-talk among these modules changes significantly from one species to another.

Gene Duplication Shapes the Functional Fate of Paralogs

Gene duplication has long been known as a major source of evolutionary novelty (*Arabidopsis* Interactome Mapping Consortium, 2011). Previous studies have found that a whole-genome duplication (WGD) event leads to more functional redundancy between paralogous proteins than small-scale duplications (SSDs) (*Arabidopsis* Interactome Mapping Consortium, 2011; Hakes et al., 2007). However, there has been much debate in the literature regarding the relative extents of sub-functionalization and neo-functionalization for diverged paralogs (Gibson and Goldberg, 2009; He and Zhang, 2005). Previous studies on functional evolution of paralogs often used interaction datasets from the literature, which, as mentioned earlier, suffer from detection and completeness biases (Yu et al., 2008). Until now, it has not been possible to measure the extent of sub-functionalization and neo-functionalization using

(G) FissionNet subnetworks of Srrm1, SPAC30D11.14C, and SPAC1952.06C.

(H) Global splicing profiles of deletion strains relative to wild-type. Columns represent total mRNA (T), pre-mRNA (P), and mature mRNA (M). Data are shown as measurements + SE. * denotes significant ($p < 0.05$); n.s. denotes not significant.

See also Figure S3 and Table S3.

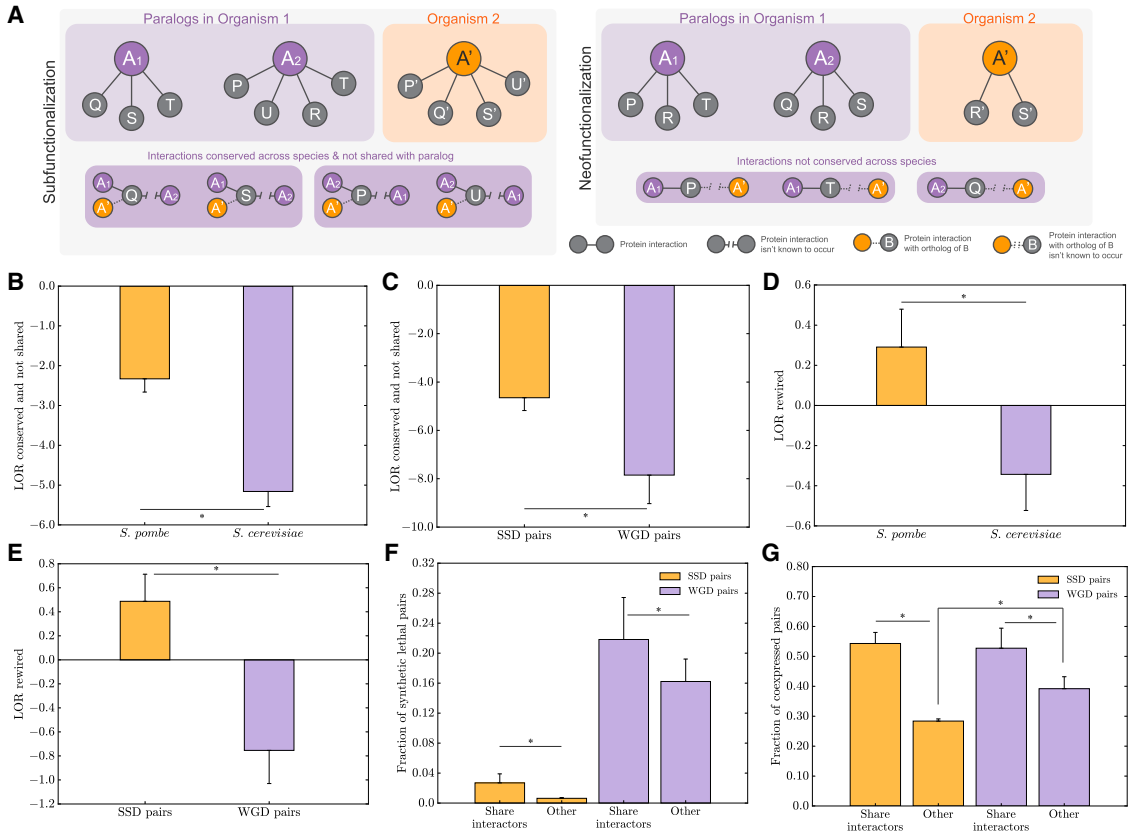


Figure 5. Functional Divergence of Interactions Involving Paralogous Proteins

(A) Schematic representation of sub-functionalization and neo-functionalization.

(B and C) Log odds ratios of sub-functionalization (B) for *S. pombe* and *S. cerevisiae* paralog pairs and (C) for *S. cerevisiae* SSD and WGD paralog pairs after correcting for divergence times.

(D and E) Log odds ratios of neo-functionalization (D) for *S. pombe* and *S. cerevisiae* paralog pairs and (E) for *S. cerevisiae* SSD and WGD paralog pairs after correcting for divergence times.

(F) Fraction of synthetic lethal pairs among SSD and WGD paralogs known or not known to share interactors.

(G) Fraction of coexpressed pairs (*PCC* > 0.4) among SSD and WGD paralogs known or not known to share interactors. Data are shown as measurements + SE.

* denotes significant ($p < 0.05$); n.s. denotes not significant.

See also Figure S5.

an unbiased framework because there was only one proteome-wide high-coverage binary protein interactome available, that of *S. cerevisiae*. Here, we compare the unbiased proteome-wide networks of *S. pombe* (FissionNet) and *S. cerevisiae* (CCSB-YI1) (Yu et al., 2008) that we produced using the same Y2H assay to analyze these two types of functional divergence.

We first examined the extent to which interactions in *S. pombe* and *S. cerevisiae* tend to be conserved across species, but not shared between within-species paralogs (sub-functionalized) (Figure 5A; see Supplemental Experimental Procedures). We find that fission yeast paralog pairs tend to undergo more sub-functionalization than budding yeast paralog pairs (Figure 5B; difference in log odds ratio = 2.8 using 1,762 fission yeast paralog pairs and 2,068 budding yeast paralog pairs, $p < 10^{-5}$, *Z* test). Since *S. pombe* paralogs arose via SSDs, while many *S. cerevisiae* paralogs arose via a WGD event (Kellis et al., 2004), this result suggests that duplication modes could impact paralog divergence differently. To test this, we compared paralog pairs generated via the WGD event with those generated via SSDs in *S. cerevisiae*. We find that SSD pairs are more sub-functionalized than WGD pairs (Figures 5C and S5A–S5D; $p < 0.05$, *Z* test; see Supplemental Experimental Procedures).

Next, we compared the extent of neo-functionalization (rewiring) (Figure 5A) for the two species and found that fission yeast paralog pairs tend to undergo more neo-functionalization than budding yeast pairs (Figures 5D and S5E; difference in log odds ratio = 0.5 using 1,158 fission yeast paralog pairs and 1,175 budding yeast paralog pairs, $p = 0.015$, *Z* test). Furthermore, within *S. cerevisiae*, SSD pairs are significantly more neo-functionalized than WGD pairs (Figures 5E and S5F–S5I; $p < 0.05$, *Z* test).

In a WGD, the entire genome is duplicated almost at once. Soon afterward, a vast majority of the duplicates are purged while only a few are retained (Kellis et al., 2004). However, the duplicates that remain are under strong evolutionary pressure to maintain stoichiometric ratios with their interaction partners

and, thus, evolve more slowly (Fares et al., 2013). On the other hand, SSDs arise sporadically and are under less pressure to maintain stoichiometric ratios, which explains why they can undergo greater functional divergence (Fares et al., 2013). This increased pressure on WGD genes to maintain stoichiometry is illustrated by their propensity to be enriched in protein complexes compared to SSD genes (Hakes et al., 2007). Using 408 high-quality literature-curated complexes from CYC2008 (Pu et al., 2009), we observed the same enrichment. Moreover, we find that the enrichment increases with the size of the complex, further supporting the notion that stoichiometric constraint influences the fate of WGD genes (Figure S5J).

Since WGD pairs are more functionally redundant than SSD pairs, these genes tend to be non-essential (Guan et al., 2007). It has also been shown that double deletions of these WGD pairs lead to a higher synthetic lethality rate than SSD pairs (Guan et al., 2007). Using a genome-scale genetic interaction map (Costanzo et al., 2010), we confirmed that deletion of WGD pairs is more likely to lead to synthetic lethality (Figure 5F; >6-fold difference in the fraction of synthetically lethal pairs, $p < 10^{-10}$, *Z* test). Moreover, when we further stratify both groups of paralogs into pairs that are known to share interactors and pairs that have not been reported to share interactors, double deletions of the former are more likely to cause synthetic lethality than double deletions of the latter (Figure 5F; ~1.5-fold difference between the two sets for both SSD and WGD pairs, $p < 0.05$ for both SSD and WGD pairs, *Z* test). This shows that paralog pairs that share interactors are more likely to be functionally redundant, regardless of whether they arose via SSD or WGD.

There have been conflicting reports in the literature regarding coexpression patterns of SSD and WGD pairs (Conant and Wolfe, 2006; Guan et al., 2007). Using a compendium of genome-wide expression datasets for *S. cerevisiae* genes (Yu et al., 2008), we found no significant difference in coexpression patterns of these pairs (Figure S5K). However, we find that SSD and WGD paralog pairs that share interactors are significantly more likely to be coexpressed than pairs that are not known to share interactors (Figures 5G, S5L, and S5M; >10% more coexpressed for paralogs that are known to share interactors, $p < 0.02$ in both cases, *Z* test; see Supplemental Experimental Procedures). The tendency to be coexpressed among SSD pairs and WGD pairs that share interactors is the same. Furthermore, among pairs that are not known to share interactors, WGD pairs tend to be more coexpressed than SSD pairs (Figures 5G, S5L, and S5M; >10% more coexpressed for WGD paralogs compared to SSD paralogs, $p < 0.02$ in all cases, *Z* test). These results show that for both duplication modes, because paralog pairs that are known to share interactors tend to be functionally redundant, the regulation of their gene expression also tends to be retained. Only for paralog pairs that are not known to share interactors is there a significant difference in coexpression between SSD and WGD paralogs, suggesting that even these WGD pairs might still be more functionally redundant than SSD pairs. It should be noted that, due to the incompleteness of current interactomes, paralog pairs could share interactors that are currently unreported.

The availability of proteome-wide interactomes helps dissect functional redundancy and divergence and to some degree the regulation of expression between paralogs. Overall, our results show that a WGD leads to greater functional redundancy while SSDs lead to greater functional diversification by sub-functionalization and neo-functionalization. Moreover, while there has been debate in the literature regarding the ubiquity of neo-functionalization (Gibson and Goldberg, 2009; He and Zhang, 2005), our results provide accurate measurements of the extent of neo-functionalization in the two yeasts.

Coevolution of Conserved Interactions Revealed by Cross-Species Interactome Mapping

To further dissect the nature of conserved interactions, we implemented a cross-species interactome mapping approach to determine the prevalence of coevolution. We consider an interaction to be coevolved when its proteins have evolved in a coordinated manner to maintain the interaction in different species but have developed incompatible binding interfaces with orthologs of their partners. To determine whether conserved interactions are intact or coevolved, we test by Y2H whether a protein in one species can interact with the ortholog of its interacting partner in another species. If the cross-species interaction can occur, the interaction is intact (Figure 4C); otherwise, it is coevolved between the two species (Figure 6A). For example, through our cross-species mapping, we discovered that interactions of farnesyltransferase subunit Cwp1 with other subunits Cpp1 and Cwg2 have coevolved between *S. pombe* and *S. cerevisiae*; Cwp1 cannot interact with either Ram1 or Cdc43, *S. cerevisiae* orthologs of Cpp1 and Cwg2, respectively (Figure 6B). A previous study showed that expression of Cwp1 cannot complement a non-functional mutant of its *S. cerevisiae* ortholog, Ram2 (Arellano et al., 1998). This suggests that Cwp1, although conserved between *S. pombe* and *S. cerevisiae* at the gene level, has evolved incompatible interaction interfaces with other farnesyltransferase subunits in *S. cerevisiae* and is thus unable to reconstitute an active enzyme complex.

It is known that evolution in protein folds is essentially the result of many random mutation events (Lockless and Ranganathan, 1999). However, since only a small fraction of changes that occur via random drift will satisfy the pairwise constraints necessary for interaction conservation, coevolution at the residue level only occurs at a few specific sites and is relatively rare (Talavera et al., 2015). Surprisingly we find that coevolution at the interaction level is not uncommon: ~33% and 50% of conserved interactions between *S. pombe* and *S. cerevisiae* or human are coevolved, respectively (Figure 6C). This shows that even among conserved interactions, only a few key alterations at important binding sites can make the cross-species interactions incompatible and the interactions coevolved. Thus, these sites are critical to protein binding and subsequent function, and changes at these sites alter protein interactions in a manner analogous to a single amino acid change disrupting protein interactions in human disease (Wang et al., 2012; Wei et al., 2014).

Among interactions for which we were able to determine coevolution status, we found that the likelihood for an interaction to be intact between *S. pombe*-*S. cerevisiae* and *S. pombe*-human is significantly higher than random expectation, while the likelihood for an interaction to be intact for one species pair

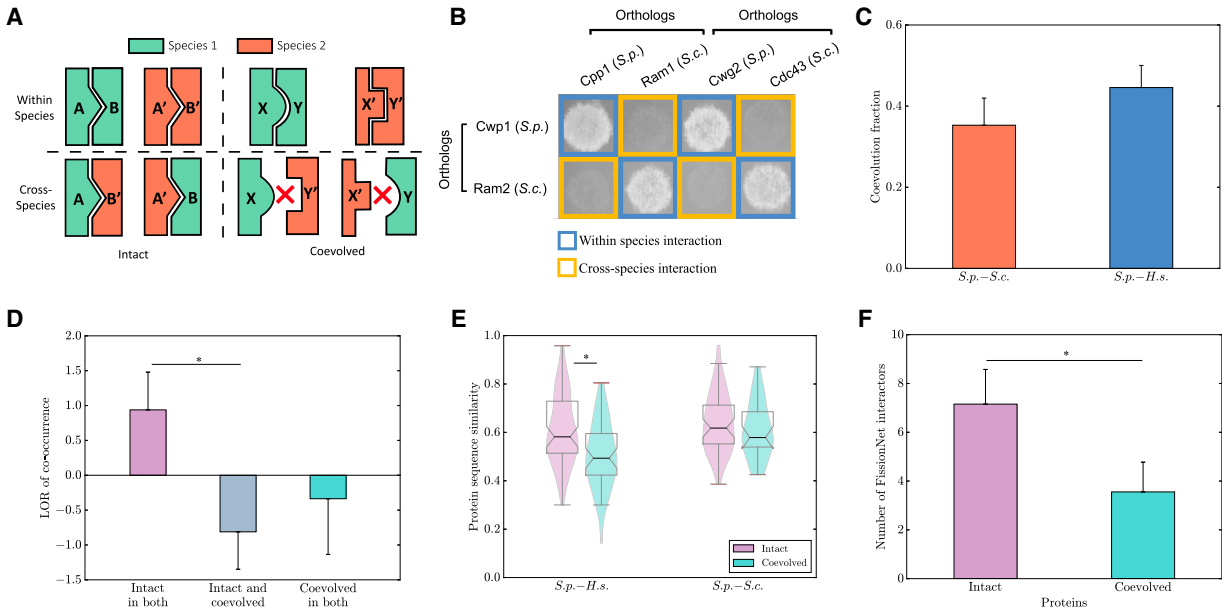


Figure 6. Intact and Coevolved Interactions

(A) Schematic representation of conserved protein interactions that are either intact or coevolved.

(B) Within- and cross-species Y2H detects coevolved interactions.

(C) Fraction of *S.p*. interactions that are coevolved with respect to *S.c*. or human (*H.s*.).

(D) Log odds ratio of co-occurrence of intact and coevolved interactions between *S.p*.-*S.c*. and *S.p*.-*H.s*.

(E) Overall protein sequence similarity of *S.p*. proteins involved in intact or coevolved interactions.

(F) Number of interactors for proteins involved in intact or coevolved interactions. Data are shown as measurements + SE. * denotes significant ($p < 0.05$); n.s. denotes not significant.

See also Figure S6 and Table S6.

and coevolved for the other species pair is significantly lower (Figure 6D; difference in log odds ratio = 1.7, $p = 0.022$, *Z* test; see Supplemental Experimental Procedures). Thus, these intact interactions are likely involved in functions that have remained unchanged among yeasts and human throughout evolution.

We then investigated potential factors that could determine whether an interaction is intact or coevolved with respect to another species. We find that overall sequences of proteins involved in intact interactions tend to be better conserved across species than sequences of proteins in coevolved interactions (Figure 6E; difference in median sequence similarity = 8.8%, $p = 2.1 \times 10^{-4}$ for *S. pombe*-human, *U* test). High sequence conservation may indicate higher levels of evolutionary constraint existing within the local network neighborhood of a given interaction. In fact, we find that proteins involved only in intact interactions have twice the number of interactors as compared to proteins involved in only coevolved interactions (Figure 6F; $p = 1.1 \times 10^{-3}$, *U* test), suggesting that the added evolutionary constraint of maintaining many interacting partners may prevent the coevolution of two interacting proteins. Finally, we find that the most highly evolutionarily correlated inter-protein residue pairs in coevolved interactions are significantly more correlated than top residue pairs in intact interactions (Figure S6A; $p < 10^{-10}$, *U* test; see Supplemental Experimental Procedures), suggesting that the maintenance of coevolved interactions involves compensatory changes at the amino acid residue level.

Implications of FissionNet for the Study of Human Disease

We explored the relevance of FissionNet to human disease by considering the context of known human disease mutations from the Human Gene Mutation Database (HGMD) (Stenson et al., 2014) within proteins of the human interactome conserved in *S. pombe*. We find that among human interactions conserved in either *S. pombe*, *S. cerevisiae*, or both, ~40% of inter-protein pairs of disease mutations cause the same disease (Figure 7A). This is significantly higher than in human interactions that are not reported to be conserved in either yeast or cannot be conserved in either due to lack of protein orthologs (Figure 7A; $p < 10^{-10}$ for all pairwise comparisons, *Z* test). Based on these results, mutations that break specific protein-protein interactions to cause diseases may be overrepresented among interactions conserved in model organisms. From a global network view, FissionNet may be highly relevant to the study of human disease based on the large portion of *S. pombe* interactions in which both proteins have human orthologs with known germline disease or somatic cancer-associated mutations (Figure 7B; 902 interactions) (Forbes et al., 2015; Stenson et al., 2014).

To demonstrate the plausibility of studying specific human disease mutations using FissionNet, we explored whether human disease mutations that disrupt human interactions intact in *S. pombe* also disrupt the corresponding interactions of the fission yeast orthologs. We focused on three examples: two

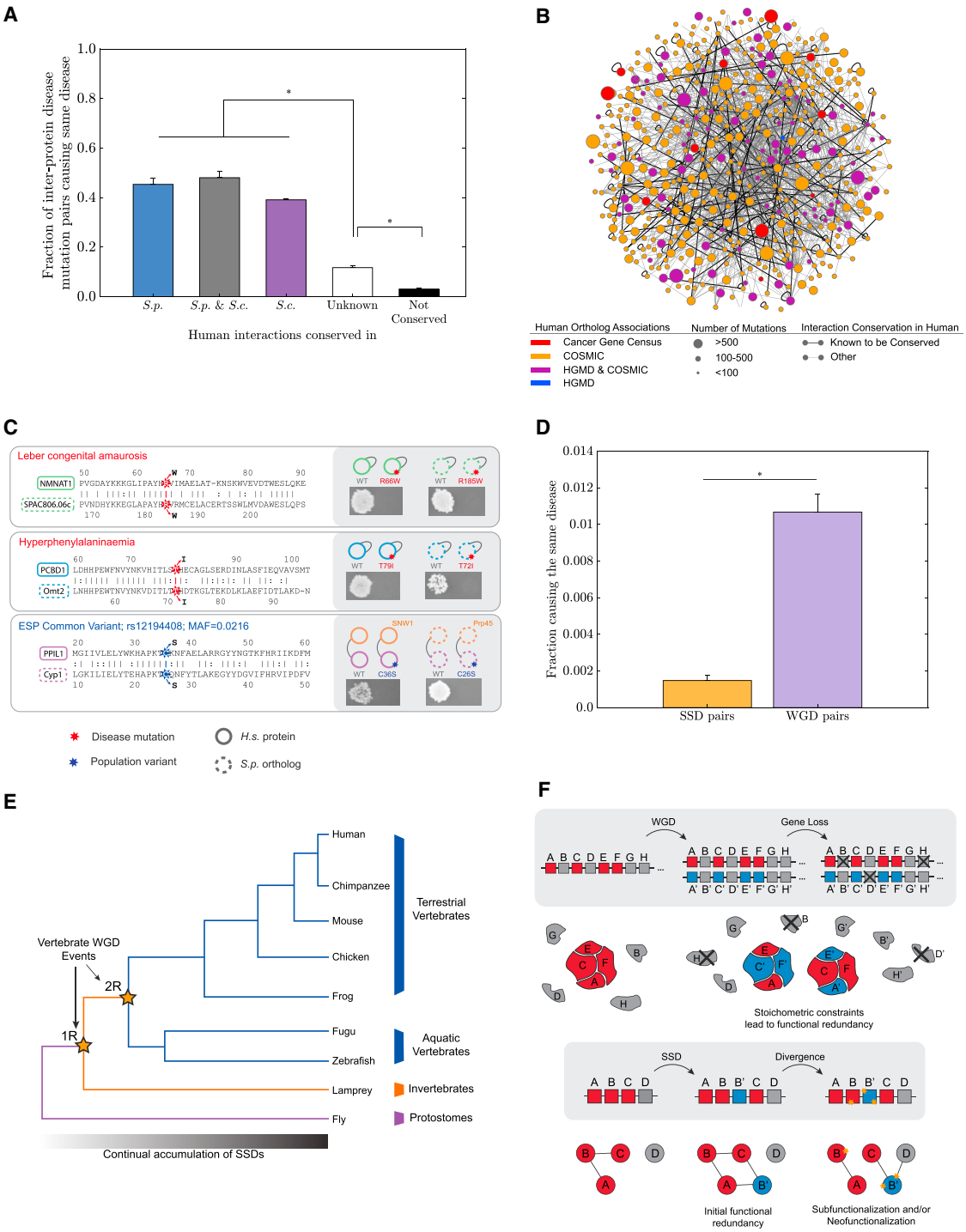


Figure 7. FissionNet as a Resource for Studying Human Disease

(A) Fraction of inter-protein HGMD mutation pairs that cause the same disease in human interactions with regard to their conservation status in *S. pombe* and *S. cerevisiae*.

(B) Largest connected subcomponent of FissionNet wherein all proteins have human orthologs with known germline disease or somatic cancer-associated mutations.

(C) Impact of human disease mutations and a population variant on intact interactions between human and fission yeast.

(D) Fraction of human SSD and WGD paralogs that cause the same disease.

(legend continued on next page)

Mendelian disease variants (Stenson et al., 2014) that disrupt the human NMNAT1-NMNAT1 and PCBD1-PCBD1 interactions and one population variant from the Exome Sequencing project (Fu et al., 2013) that disrupts the human SNW1-PPIL1 interaction. We find that introducing these human protein residue changes into their *S. pombe* orthologs also disrupts the fission yeast interactions (Figure 7C). These results indicate that cross-species interactome mapping enables investigation of whether interaction interfaces are altered at the molecular level between model organisms and human, a finding with potentially far-reaching implications for the study of protein function and human disease.

Our results regarding gene duplication modes may also be relevant to the study of human disease. We find that human WGD paralog pairs have a significantly higher likelihood to be involved in the same disease compared to human SSD paralog pairs, in agreement with our observation that WGD paralog pairs tend to be functionally redundant (Figure 7D; 7-fold difference in the fraction of WGD and SSD pairs that cause the same disease, $p < 10^{-10}$, *Z* test; see Supplemental Experimental Procedures). Thus, our findings have direct implications for understanding the functional roles of paralogous genes, from yeasts to human.

DISCUSSION

FissionNet provides a wealth of functional information. For example, we find that the Atf1-Cid12 interaction mediates silencing at Atf1-target genes *hsp16* and *hsp104*. It has been shown that the RNAi pathway is involved in silencing of these genes (Woolcock et al., 2012). Hence, it is possible that the Atf1-Cid12 interaction is part of an RNAi-dependent regulatory pathway.

By comparing FissionNet to protein networks in budding yeast and human, we have shown that the molecular bases for interaction conservation among orthologous proteins are complex and different from those that underlie gene conservation. This is highly relevant to the use of the two yeasts as model organisms as there are functions that can be better studied using fission yeast. We find that divergence across species is not completely dictated by sequence level changes, suggesting that rewiring of interactomes plays an important role in species evolution. Additionally, our finding that proteins in a significant fraction of conserved interactions have undergone coevolution to maintain interactions has major implications for studies reliant on the expression of human proteins in model organisms to identify functional mechanisms (Tardiff et al., 2013).

Gene Duplications Introduce Evolutionary Innovation and Robustness

Gene duplication is a key process shaping evolution (Figure 7E). Our results show that paralogs arising via WGD are under strong constraints to maintain stoichiometric ratios with their interaction partners and, hence, tend to maintain functional redundancy; on the other hand, duplicates arising via SSDs are not under such strong constraints and are more likely to gain novel functions (Figure 7F). For example, it has previously been reported that duplicate copies of the *SRGAP2* gene that arose via segmental duplications (SSD-like events) have gained new functions related to brain development specifically in the human lineage (Dennis et al., 2012).

Gene duplications play an important role in the evolutionary mechanism governing speciation as well as the evolution of developmental and morphological complexity in vertebrates (Rensing, 2014; Ting et al., 2004). For example, two rounds of WGD have been predicted in the origin of jawed vertebrates (Figure 7E) (Kasahara, 2007). During speciation, while certain key functions need to be evolutionarily preserved, new functions are necessary for differential adaptation between species (Ting et al., 2004). Previous studies have identified how duplication events can lead to functional changes through gene dosage alterations (Papp et al., 2003). Our results help establish on a proteomic scale that paralogs arising via WGD are more likely to preserve functions and provide robustness for important cellular functions, while paralogs arising via SSDs are more likely to contribute to novel functions gained by specific species. These findings further our understanding of human biology and disease.

Future Directions

Our analyses focus on budding yeast, fission yeast, and human, as they are the only three eukaryotic organisms for which we have proteome-scale interactome networks using our version of the Y2H assay (>50% of all protein pairs screened). When more interactome networks are systematically generated in other species, using assays with measured sensitivity and specificity, the comparative network analysis framework established in this study can be readily applied to further elucidate the extent and nature of the evolution of protein functions across many species.

EXPERIMENTAL PROCEDURES

Generation of the Binary Protein-Protein Interactome Map of *S. pombe*

FissionNet was generated by triplicate independent screening of ~4,900 *S. pombe* open reading frame (ORFs) (Figure S1A). The network was validated by testing a representative 220 interacting ORF pairs using PCA assays and by determining its functional properties with respect to random pairs and to a literature-curated network.

Conservation of Interactions in *S. pombe*, *S. cerevisiae*, and Human

We focused only on interactions that can be conserved, i.e., both proteins involved in the interaction have orthologs in the other species. We mapped interactions in the reference species to their corresponding ortholog pairs in the other species and tested these pairs using our Y2H assay in a pairwise fashion. Overall, results from these pairwise retests for all three species (a total of ~20,000 individual Y2H experiments) are used to obtain the observed conservation fraction. To accurately estimate the true conservation fraction, we developed a rigorous Bayesian framework that takes into account both the false-positive and false-negative rates of our Y2H assay and computes the true conservation fraction from the observed fraction.

(E) The 2R hypothesis predicts two recent WGD events leading to the vertebrate lineage.

(F) WGD can lead to more functional redundancy through targeted gene loss that maintains stoichiometric ratios of protein products. SSD leads to more neofunctionalization and sub-functionalization through alterations to initially redundant paralogs. Data are shown as measurements + SE. * denotes significant ($p < 0.05$); n.s. denotes not significant.

Positive and Negative Reference Sets

The PRS and NRS constitute sets of positive and negative controls, respectively. Our PRS comprises 93 *S. pombe* interactions that have been previously reported in two or more publications. To construct the NRS we chose 168 random protein pairs that have not been reported to interact in *S. pombe* and whose orthologs have not been reported to interact in any species. In a set of random protein pairs, the expected fraction of interactions is $\sim 10^{-3}$–10^{-4} (Riley et al., 2005; Yu et al., 2008), the expected number of interactions in a random set of 168 pairs is $< 10^{-3} \times 168$ (≈ 0.2). Since we exclude pairs that are known to interact, the expected number of interactions in our NRS is even lower.

See the Supplemental Experimental Procedures for a full description of the materials and methods.

ACCESSION NUMBERS

The accession number for the DNA microarray data reported in this paper is NCBI GEO: GSE74354.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2015.11.037.

AUTHOR CONTRIBUTIONS

H. Yu conceived and oversaw all aspects of the project. T.V.V. and H. Yu designed experiments. J.D. performed computational analyses pertaining to gene duplication. J.D. and M.J.M. performed all other computational analyses. T.V.V., N.A.C., N.A., A.G.D., R.F., and L.G.L. performed Y2H experiments with computational support from J.D. and M.J.M. N.A. performed PCA experiments. B.J.F. and J.A.P. performed DNA microarray experiments, and T.V.V., B.J.F., and M.J.M. interpreted the data. T.V.V. and X.W. performed western blotting and coimmunoprecipitation experiments. T.V.V. and S.H. performed RT-PCR experiments. T.V.V. performed all other experiments. A.M., M.T., H. Yamano, and M.Y. generated the *S. pombe* entry clone ORFeome libraries. A.G., F.P.R., J.A.P., Y.X., and H. Yu supervised research. T.V.V., J.D., M.J.M., and H. Yu wrote the manuscript with input from other coauthors.

ACKNOWLEDGMENTS

We thank D. Moazed for *S. pombe* strains provided for the analyses of centromeric silencing; S. Forsburg, H. Madhani, and J. Nakayama for providing plasmids for protein expression in *S. pombe*. We are grateful to M. Smolka for experimental advice; D. Barbash, A. Clark, S. Emr, S. Grewal, D. Lin, D. Shalloway, and members of the Yu lab for discussions. The accession number for the DNA microarray data reported in this paper is NCBI GEO: GSE74354. This work was supported by Marie Curie Cancer Care, Association for International Cancer Research, and Cancer Research UK to M.T. and H. Yamano; NHGRI grants HG001715 and HG004233, Krembil Foundation, and Canada Excellence Research Chairs Program to F.P.R.; Canada Research Chairs Program and grants NSERC RGPIN-2014-03892, NSF CCF-1219007, and CFI JELF-33732 to Y.X.; and NIGMS grants GM098634 to J.A.P. and GM097358 and GM104424 to H. Yu.

Received: July 2, 2015
Revised: October 12, 2015
Accepted: November 4, 2015
Published: January 14, 2016

REFERENCES

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

Arellano, M., Coll, P.M., Yang, W., Duran, A., Tamanoi, F., and Perez, P. (1998). Characterization of the geranylgeranyl transferase type I from Schizosaccharomyces pombe. Mol. Microbiol. *29*, 1357–1367.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al.; The Gene Ontology Consortium (2000). Gene ontology: tool for the unification of biology. Nat. Genet. *25*, 25–29.

Bader, J.S., Chaudhuri, A., Rothberg, J.M., and Chant, J. (2004). Gaining confidence in high-throughput protein interaction networks. Nat. Biotechnol. *22*, 78–85.

Blencowe, B.J., Issner, R., Nickerson, J.A., and Sharp, P.A. (1998). A coactivator of pre-mRNA splicing. Genes Dev. *12*, 996–1009.

Conant, G.C., and Wolfe, K.H. (2006). Functional partitioning of yeast coexpression networks after genome duplication. PLoS Biol. *4*, e109.

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E.D., Sevier, C.S., Ding, H., Koh, J.L., Toufighi, K., Mostafavi, S., et al. (2010). The genetic landscape of a cell. Science *327*, 425–431.

Das, J., and Yu, H. (2012). HINT: High-quality protein interactomes and their applications in understanding human disease. BMC Syst. Biol. *6*, 92.

Das, J., Vo, T.V., Wei, X., Mellor, J.C., Tong, V., Degatano, A.G., Wang, X., Wang, L., Cordero, N.A., Kruer-Zerhusen, N., et al. (2013). Cross-species protein interactome mapping reveals species-specific wiring of stress response pathways. Sci. Signal. *6*, ra38.

Dennis, M.Y., Nuttle, X., Sudmant, P.H., Antonacci, F., Graves, T.A., Nefedov, M., Rosenfeld, J.A., Sajjadian, S., Malig, M., Kotkiewicz, H., et al. (2012). Evolution of human-specific neural SRGAP2 genes by incomplete segmental duplication. Cell *149*, 912–922.

Fares, M.A., Keane, O.M., Toft, C., Carretero-Paulet, L., and Jones, G.W. (2013). The roles of whole-genome and small-scale duplications in the functional specialization of Saccharomyces cerevisiae genes. PLoS Genet. *9*, e1003176.

Forbes, S.A., Beare, D., Gunasekaran, P., Leung, K., Bindal, N., Boutselakis, H., Ding, M., Bamford, S., Cole, C., Ward, S., et al. (2015). COSMIC: exploring the world's knowledge of somatic mutations in human cancer. Nucleic Acids Res. *43*, D805–D811.

Fu, W., O'Connor, T.D., Jun, G., Kang, H.M., Abecasis, G., Leal, S.M., Gabriel, S., Rieder, M.J., Altshuler, D., Shendure, J., et al.; NHLBI Exome Sequencing Project (2013). Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature *493*, 216–220.

Gibson, T.A., and Goldberg, D.S. (2009). Questioning the ubiquity of neofunctionalization. PLoS Comput. Biol. *5*, e1000252.

Giot, L., Bader, J.S., Brouwer, C., Chaudhuri, A., Kuang, B., Li, Y., Hao, Y.L., Ooi, C.E., Godwin, B., Vitols, E., et al. (2003). A protein interaction map of Drosophila melanogaster. Science *302*, 1727–1736.

Guan, Y., Dunham, M.J., and Troyanskaya, O.G. (2007). Functional analysis of gene duplications in Saccharomyces cerevisiae. Genetics *175*, 933–943.

Hakes, L., Pinney, J.W., Lovell, S.C., Oliver, S.G., and Robertson, D.L. (2007). All duplicates are not equal: the difference between small-scale and genome duplication. Genome Biol. *8*, R209.

He, X., and Zhang, J. (2005). Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics *169*, 1157–1164.

Hegele, A., Kamburov, A., Grossmann, A., Sourlis, C., Wowro, S., Weimann, M., Will, C.L., Pena, V., Lührmann, R., and Stelzl, U. (2012). Dynamic protein-protein interaction wiring of the human spliceosome. Mol. Cell *45*, 567–580.

Holoch, D., and Moazed, D. (2015). RNA-mediated epigenetic regulation of gene expression. Nat. Rev. Genet. *16*, 71–84.

Johnson, A.E., Chen, J.S., and Gould, K.L. (2013). CK1 is required for a mitotic checkpoint that delays cytokinesis. Curr. Biol. *23*, 1920–1926.

Kamada, K., Shu, F., Chen, H., Malik, S., Stelzer, G., Roeder, R.G., Meisterernst, M., and Burley, S.K. (2001). Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription complex. Cell *106*, 71–81.

Kasahara, M. (2007). The 2R hypothesis: an update. Curr. Opin. Immunol. *19*, 547–552.

Kellis, M., Birren, B.W., and Lander, E.S. (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature *428*, 617–624.

Lockless, S.W., and Ranganathan, R. (1999). Evolutionarily conserved pathways of energetic connectivity in protein families. Science *286*, 295–299.

Matthews, L.R., Vaglio, P., Reboul, J., Ge, H., Davis, B.P., Garrels, J., Vincent, S., and Vidal, M. (2001). Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or "interologs". Genome Res. *11*, 2120–2126.

Motamedi, M.R., Verdel, A., Colmenares, S.U., Gerber, S.A., Gygi, S.P., and Moazed, D. (2004). Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. Cell *119*, 789–802.

Nepusz, T., Yu, H., and Paccanaro, A. (2012). Detecting overlapping protein complexes in protein-protein interaction networks. Nat. Methods *9*, 471–472.

Papp, B., Pál, C., and Hurst, L.D. (2003). Dosage sensitivity and the evolution of gene families in yeast. Nature *424*, 194–197.

Pu, S., Wong, J., Turner, B., Cho, E., and Wodak, S.J. (2009). Up-to-date catalogues of yeast protein complexes. Nucleic Acids Res. *37*, 825–831.

Rensing, S.A. (2014). Gene duplication as a driver of plant morphogenetic evolution. Curr. Opin. Plant Biol. *17*, 43–48.

Riley, R., Lee, C., Sabatti, C., and Eisenberg, D. (2005). Inferring protein domain interactions from databases of interacting proteins. Genome Biol. *6*, R89.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J., et al. (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. Science *322*, 405–410.

Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S.J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., Fontanillo, C., Mosca, R., et al. (2014). A proteome-scale map of the human interactome network. Cell *159*, 1212–1226.

Sahni, N., Yi, S., Taipale, M., Fuxman Bass, J.I., Coulombe-Huntington, J., Yang, F., Peng, J., Weile, J., Karras, G.I., Wang, Y., et al. (2015). Widespread macromolecular interaction perturbations in human genetic disorders. Cell *161*, 647–660.

Shiozaki, K., and Russell, P. (1996). Conjugation, meiosis, and the osmotic stress response are regulated by Spc1 kinase through Atf1 transcription factor in fission yeast. Genes Dev. *10*, 2276–2288.

Sipiczki, M. (2000). Where does fission yeast sit on the tree of life? Genome Biol. *1*, reviews1011.1–reviews1011.4.

Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., et al. (2005). A human protein-protein interaction network: a resource for annotating the proteome. Cell *122*, 957–968.

Stenson, P.D., Mort, M., Ball, E.V., Shaw, K., Phillips, A., and Cooper, D.N. (2014). The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. Hum. Genet. *133*, 1–9.

Talavera, D., Lovell, S.C., and Whelan, S. (2015). Covariation Is a Poor Measure of Molecular Coevolution. Mol. Biol. Evol. *32*, 2456–2468.

Tardiff, D.F., Jui, N.T., Khurana, V., Tambe, M.A., Thompson, M.L., Chung, C.Y., Kamadurai, H.B., Kim, H.T., Lancaster, A.K., Caldwell, K.A., et al. (2013). Yeast reveal a "druggable" Rsp5/Nedd4 network that ameliorates α-synuclein toxicity in neurons. Science *342*, 979–983.

Ting, C.T., Tsaur, S.C., Sun, S., Browne, W.E., Chen, Y.C., Patel, N.H., and Wu, C.I. (2004). Gene duplication and speciation in Drosophila: evidence from the Odysseus locus. Proc. Natl. Acad. Sci. USA *101*, 12232–12235.

Verdel, A., Jia, S., Gerber, S., Sugiyama, T., Gygi, S., Grewal, S.I., and Moazed, D. (2004). RNAi-mediated targeting of heterochromatin by the RITS complex. Science *303*, 672–676.

Wang, X., Wei, X., Thijssen, B., Das, J., Lipkin, S.M., and Yu, H. (2012). Three-dimensional reconstruction of protein networks provides insight into human genetic disease. Nat. Biotechnol. *30*, 159–164.

Wei, X., Das, J., Fragoza, R., Liang, J., Bastos de Oliveira, F.M., Lee, H.R., Wang, X., Mort, M., Stenson, P.D., Cooper, D.N., et al. (2014). A massively parallel pipeline to clone DNA variants and examine molecular phenotypes of human disease mutations. PLoS Genet. *10*, e1004819.

Wood, V., Gwilliam, R., Rajandream, M.A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., et al. (2002). The genome sequence of Schizosaccharomyces pombe. Nature *415*, 871–880.

Woolcock, K.J., Stunnenberg, R., Gaidatzis, D., Hotz, H.R., Emmerth, S., Barraud, P., and Bühler, M. (2012). RNAi keeps Atf1-bound stress response genes in check at nuclear pores. Genes Dev. *26*, 683–692.

Yu, H., Braun, P., Yildirim, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., et al. (2008). High-quality binary protein interaction map of the yeast interactome network. Science *322*, 104–110.

Yu, H., Tardivo, L., Tam, S., Weiner, E., Gebreab, F., Fan, C., Svrzikapa, N., Hirozane-Kishikawa, T., Rietman, E., Yang, X., et al. (2011). Next-generation sequencing to generate interactome datasets. Nat. Methods *8*, 478–480.