An in Vivo Map of the Yeast Protein Interactome

Kirill Tarassov *et al.*

Science **320**, 1465 (2008);
DOI: 10.1126/science.1153878

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

An in Vivo Map of the Yeast Protein Interactome

Kirill Tarassov,[1]* Vincent Messier,[1]* Christian R. Landry,[1,2]* Stevo Radinovic,[1]* Mercedes M. Serna Molina,[1] Igor Shames,[1] Yelena Malitskaya,[1] Jackie Vogel,[3] Howard Bussey,[3] Stephen W. Michnick[1,2]†

Protein interactions regulate the systems-level behavior of cells; thus, deciphering the structure and dynamics of protein interaction networks in their cellular context is a central goal in biology. We have performed a genome-wide in vivo screen for protein-protein interactions in *Saccharomyces cerevisiae* by means of a protein-fragment complementation assay (PCA). We identified 2770 interactions among 1124 endogenously expressed proteins. Comparison with previous studies confirmed known interactions, but most were not known, revealing a previously unexplored subspace of the yeast protein interactome. The PCA detected structural and topological relationships between proteins, providing an 8-nanometer–resolution map of dynamically interacting complexes in vivo and extended networks that provide insights into fundamental cellular processes, including cell polarization and autophagy, pathways that are evolutionarily conserved and central to both development and human health.

The elucidation of protein-protein interaction networks (PINs, or interactomes) holds the promise of answering fundamental questions about how the biochemical machinery of cells organizes matter, information, and energy transformations to perform specific functions (*1*). An essential and rarely addressed question is whether protein complexes and PINs that are reconstructed or reconstituted in vitro or removed from the normal context in which they are expressed reflect their organization in living cells. For eukaryotes, the test bed for large-scale analysis of PINs is the yeast *Saccharomyces cerevisiae,* where several PIN analyses have been performed using yeast two-hybrid screens (Y2H) (*2*–*4*) or tandem affinity purification followed by mass-spectrometric analyses (TAP-MSs) (*5*–*8*). Each approach captures specific features of protein interactions; two-hybrid methods are best at measuring direct binary interactions between pairs of proteins, whereas affinity purification techniques best capture stable protein complexes. However, neither approach measures interactions between proteins in their natural cellular context, and are not easily amenable to studying protein complexes that are transiently associated or dynamic under different conditions, that do not survive in vitro purification, or that cannot be transported to the nucleus and form interactions in the absence of other

[1]Département de Biochimie, Université de Montréal Casier postal 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7, Canada. [2]Centre Robert-Cedergren, Bio-Informatique et Génomique, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7, Canada. [3]Department of Biology, McGill University, 1205 Avenue Dr. Penfield, Montréal, Québec H3A 1B1, Canada.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: stephen.michnick@umontreal.ca

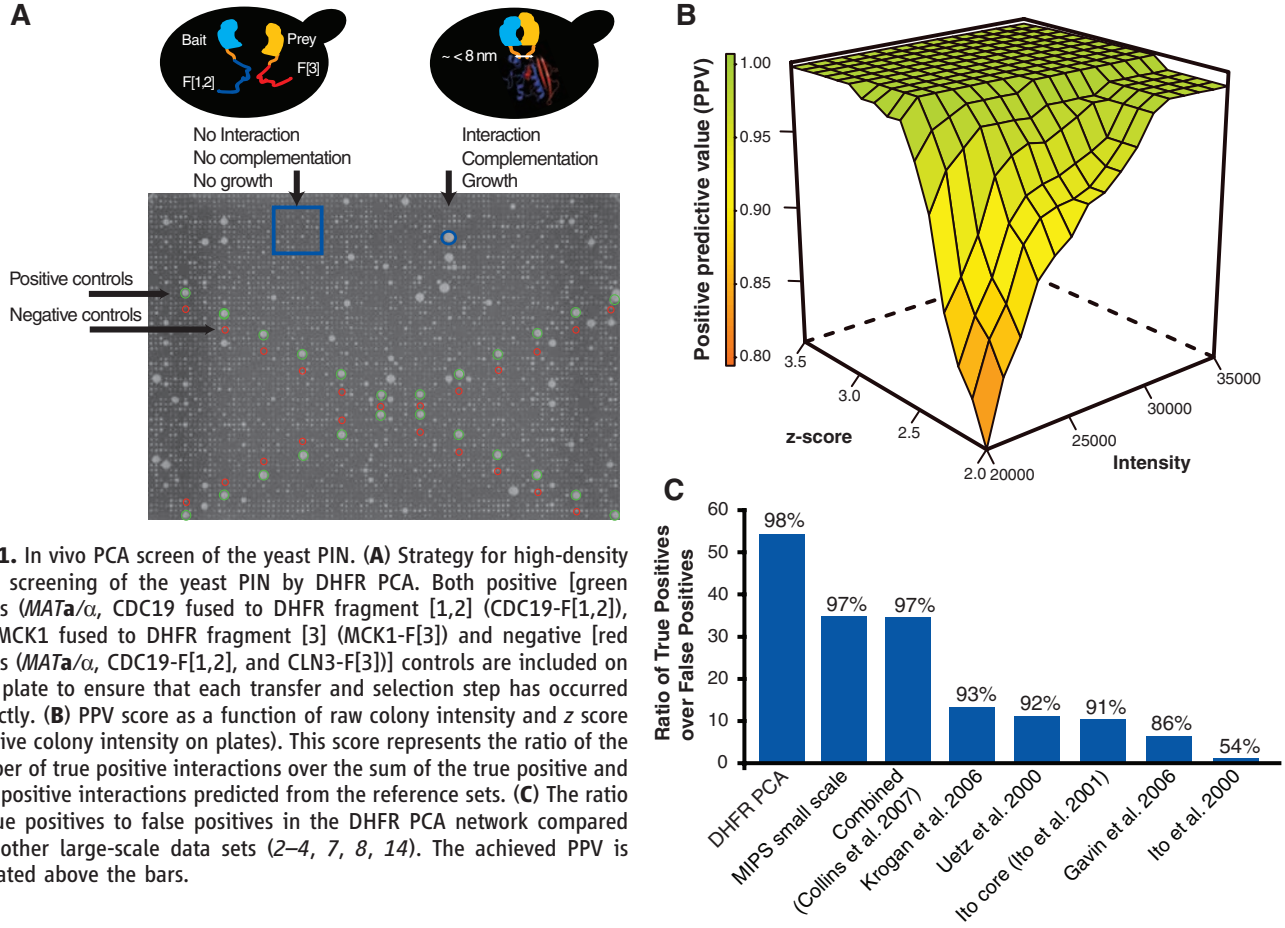


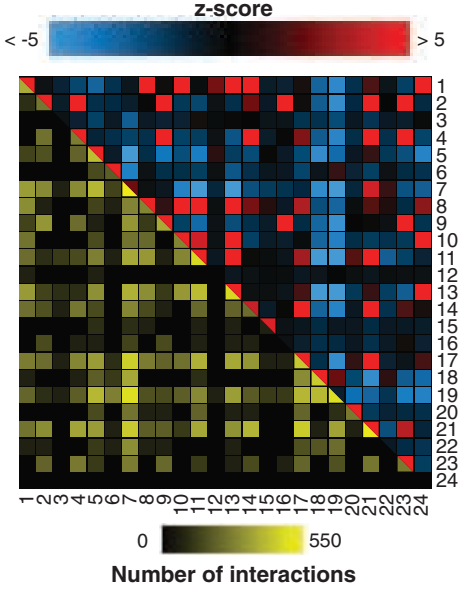
Fig. 1. In vivo PCA screen of the yeast PIN. (**A**) Strategy for high-density array screening of the yeast PIN by DHFR PCA. Both positive [green circles (*MAT***a**/α, CDC19 fused to DHFR fragment [1,2] (CDC19-F[1,2]), and MCK1 fused to DHFR fragment [3] (MCK1-F[3]) and negative [red circles (*MAT***a**/α, CDC19-F[1,2], and CLN3-F[3])] controls are included on each plate to ensure that each transfer and selection step has occurred correctly. (**B**) PPV score as a function of raw colony intensity and *z* score (relative colony intensity on plates). This score represents the ratio of the number of true positive interactions over the sum of the true positive and false positive interactions predicted from the reference sets. (**C**) The ratio of true positives to false positives in the DHFR PCA network compared with other large-scale data sets (*2*–*4*, *7*, *8*, *14*). The achieved PPV is indicated above the bars.

stabilizing interactions as necessitated in Y2H screening. Protein-fragment complementation assays (PCA) provide an alternative approach to detect protein-protein interactions (PPIs) in their natural context. In the PCA strategy, two proteins of interest are fused to complementary fragments of a reporter protein. If the proteins of interest interact physically, the reporter fragments are brought together and fold into their native structure, thus reconstituting the reporter activity of the PCA (Fig. 1A). PCA strategies provide a simple, direct means for the detection of PPIs in vivo, and do so with endogenously expressed full-length proteins in their native posttranslationally modified states and cellular locations (*9*). Further, PCAs provide spatial and topological information about PPIs. Thus, a large-scale PCA screen would provide direct insights into the global structural organization of PINs as they exist in the living cell.

Genome-wide in vivo screen. We have performed a systematic binary screen for PPIs at a genomewide scale in *S. cerevisiae* using a PCA based on the murine dihydrofolate reductase (mDHFR) assay adapted to yeast (*10*) (Fig. 1A). The DHFR PCA is a survival-selection assay based on a mutant of mDHFR that is insensitive to the DHFR inhibitor methotrexate but retains full catalytic activity and allows detection of PPIs with as few as 25 to 100 complexes per cell (*11*, *12*). We created unique homologous recombination cassettes for all 5756 consensus genes with both the F[1,2] and F[3] complementary N- and C-terminal DHFR fragment sequences (*10*). Successful cassette transformation of *S. cerevisiae* haploids was achieved for 4326 (75%) open reading frames (ORFs) with the DHFR F[1,2] fragment in *MAT***a** and 4804 (83%) ORFs with the DHFR F[3] fragment in *MAT*α strains, with a final combined coverage of 5367 (93%) of all ORFs (table S1). The entire screening process was performed on solid-phase medium (Fig. 1A and fig. S1). Briefly, *MAT***a** strains (F[1,2] fragment fusions) served as baits and were mated individually with all *MAT*α (F[3]) strains on high-density arrays. The resulting diploids were transferred to a minimal medium [synthetic complete (SC)] plate to select for methotrexate resistance (reconstituted mDHFR activity, with native *S. cerevisiae* DHFR inhibited), and colony growth was recorded using automated analysis of digital images (Fig. 1A and figs. S2 and S3). PPIs were determined based on the growth of the diploid colonies measured by the pixel intensities on the selection plates (figs. S1 to S3). In total, 3247 individual highly reproducible (fig. S4) bait screens were performed, resulting in more than 15 million individual matings.

Data filtering, quality assessment, and overlap with existing PINs. We experimentally accounted for two potential sources of false positives in a PCA screen: trapping of nonspecific complexes due to irreversible folding of the mDHFR reporter protein, and potential spontaneous complementation (folding) of the DHFR PCA fragments. First, we used the adenosine 3′,5′-monophosphate–dependent dissociation of the yeast protein kinase A complex as a test system (*13*) to show that the DHFR PCA is fully reversible, and thus the trapping of complexes is unlikely (*10*). Second, we screened all the strains against the individual F[1,2] and F[3] complementary fragments or fragment-peptide linker sequences. This allowed us to eliminate 344 promiscuous, highly expressed proteins (fig. S6 and table S2), several of which are also often observed as false-positives in affinity purifications (*10*). We next identified a threshold of colony intensity above which we could infer PPI. The Munich Information Center for Protein Sequences (MIPS) complexes were used as a standard set of true positives, along with 266,858 true negative interactions between proteins expressed in different cellular compartments or having negatively correlated expression (*14*, *15*). After several filtering steps (*10*) and benchmarking on the reference PPIs, we obtained a high-quality data set containing 2770 interactions among 1124 proteins that reach a positive predictive value (PPV) of 98.2% (Fig. 1B and tables S3 and S4). This resulted in data having precision (number of true positives relative to false positives) comparable to the MIPS small-scale experiments and all previous large-scale data sets (Fig. 1C and fig. S5). The proteins in the DHFR PCA network are highly enriched in cellular compartments [for example, organelle membranes ($P < 10^{-12}$), proteasome regulatory particles ($P < 10^{-8}$), the nucleolus ($P < 10^{-7}$), and the cell cortex ($P < 10^{-7}$)] that were less represented in comprehensive TAP-MS results (*14*) (tables S5 and S6). The high sensitivity of the DHFR PCA assay is reflected in the abundance of the proteins that populated our network, which are on average only slightly more expressed than the proteome [the median log10(protein abundance) = 2.32 versus 2.28; Wilcoxon rank sum test, $P = 0.19$] and spanning the whole distribution of protein abundance (fig. S6).

Because this study was performed in vivo, with a technique never used at this scale and in a different medium than previous experiments, we expected that many interactions would be previously undiscovered. An examination of major databases of PPIs reveals that most of the interactions (~80%) we report are among protein pairs for which no data had been previously reported (fig. S7A). However, when considering only PPIs that could be detected by both DHFR PCA and the other experiments (*10*), we confirmed between 16 and 41% of PPIs reported in previous large-scale screens, suggesting excellent concordance between the results of our and very different methods (figs. S7B and S8). Further, PPIs derived from PCA represent pairwise interactions, which contrasts with TAP-MS PINs, which identify clusters and thus complexes of interacting proteins. PPIs detected by PCA are therefore either within, between, or outside these complexes and thus complement these previous studies. For instance, 10% of the DHFR PCA PPIs map within specific complexes in the combined analyses of the two TAP-MS data sets (*7*, *15*), and 36 and 38% of the DHFR PCA PPIs are between one protein found in a complex and one protein not in the published data set, or two proteins not in the data set, respectively. We identified several inter-

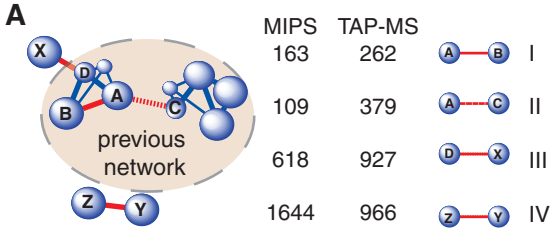


(1) Golgi apparatus; (2) cell cortex; (3) cell wall; (4) cellular bud; (5)unknown; (6) chromosome; (7) cytoplasm; (8) cytoplasmic membrane-bound vesicle; (9) cytoskeleton; (10) endomembrane system; (11) endoplasmic reticulum; (12) extracellular region; (13) membrane; (14) membrane fraction; (15) microtubule organizing center; (16) mitochondrial envelope; (17) mitochondrion; (18) nucleolus; (19) nucleus; (20) peroxisome; (21) plasma membrane; (22) ribosome; (23) site of polarized growth; (24) vacuole;

Fig. 2. Interactions are enriched within GO categories. The DHFR PCA network covers several classes of protein function, location, and biological process. The colors above the diagonal represent positive and negative deviations from the expected number of interactions between two cell compartments. A positive *z* score indicates a larger number of interactions within or between two categories as compared with a random network. A negative *z* score indicates a smaller number of interactions than expected. A *z* score of 2 or −2 corresponds to a *P* value of 0.05, and a *z* score of 5 or −5 to a *P* value of 5×10^{-7}. Values below −5 and above 5 were given these minimal and maximal values. Entries below the diagonal indicate the observed numbers of interactions on a $\log_{10}$ scale.

actions among complexes (15%), which probably mediate the integration of biological processes among PIN modules. For instance, PPIs occur among complexes that are more related in their functional annotations than would be expected to occur by chance [interacting protein pairs had a semantic similarity score of cell compartments (CCs) of 3.44 versus 1.64, $P < 10^{-100}$; of biological processes (BPs) of 3.48 versus 1.51, $P < 10^{-80}$; and of molecular functions (MFs) of 3.53 versus 2.3, $P < 10^{-10}$]. For example, we see interactions between Dhh1p and Lsm4p, both involved in the RNA metabolic process but part of the CCR4 and the RNA-splicing complex, respectively. Another example is the interaction between Reg1p and Snf1p: subunits of the serine-threonine phosphoprotein phosphatase and SNF1 complex, respectively, but both involved in the regulation of carbohydrate metabolic processes (table S7). Finally, we report 286 interactions involving one uncharacterized protein with proteins of known function ($n = 278$ interactions) or between two uncharacterized proteins ($n = 8$) (*16*), which will aid in their functional annotation.

General organization of the yeast DHFR PCA PIN. Because we detected PPIs as they occurred in intact cells, with the faithful representation of gene expression timing and protein localization, we predicted and observed stronger coregulation of interacting protein pairs (Pearson $r = 0.2$ versus $r = 0.1$, $P << 0.001$) than was expected for random networks of the same size with the same protein connectivity. This is also mirrored in the enrichment of interactions among proteins that share the same BPs, MFs, and CCs and a depletion of interactions among genes of different categories (Fig. 2 and fig. S9). PPIs among categories are somewhat more enriched in the PCA-determined network as compared with TAP-MS studies. For instance, 64, 56, and 63% of DHFR PCA interactions map to different BPs, CCs, and MFs, whereas these numbers are smaller in the TAP-MS PINs [58, 46, and 57% (*8*) and 51, 49, and 58% (*7*)]. Much of this increased enrichment of the cross-cellular components reflects interactions among proteins that the DHFR PCA method covers more of than TAP-MS; these are interactions that appear to represent the natural exchange of proteins between, for instance, the endoplasmic reticulum, Golgi, mitochondrial envelope, and vacuolar proteins, whereas others reflect the organization of complex cellular processes. For example, high enrichments in interactions between proteins localized to the bud and bud neck with those of the cell cortex, cytoskeleton, plasma membrane, and sites of polarized growth reflect the roles of these proteins in several compartments during cell division. We also saw strong compartmentalization of interactions; for example, for nuclear and nucleolar proteins, which show enrichment in interactions between proteins in these two compartments but strong depletions in interactions with those of any other compartment. Equally, patterns of cross-process and molecular function categories reflect differences in complexity and organization (fig. S9). For example, among molecular functions, RNA binding is specifically enriched in interactions between helicase and translation regulatory functions, whereas the more general transporter activity category shows links to diverse functions. The observation that PCA interactions detect links among functionally related categories is supported by a semantic analysis of the full Gene Ontology (GO) hierarchies. Proteins that show interactions with different GO Slim annotations have higher semantic similarities in their GO terms than expected by chance (CCs, 1.52 versus 0.94, $P < 10^{-231}$; BPs, 2.04 versus 1.35, $P < 10^{-122}$; and MFs, 1.89 versus 1.64, $P < 10^{-8}$), and may thus represent inter-

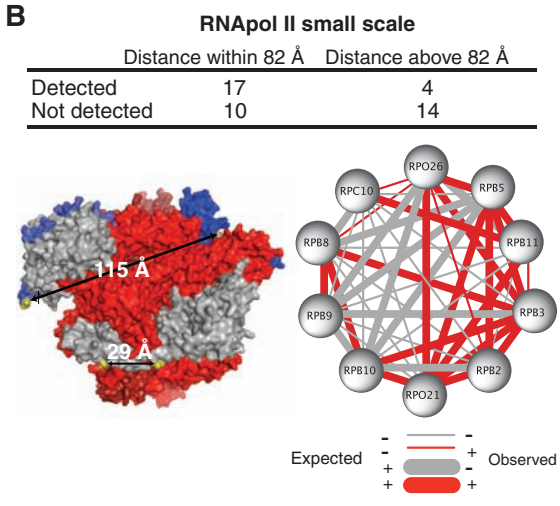


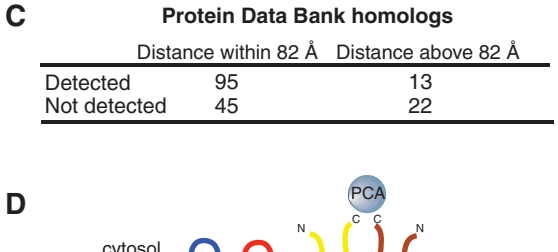
Type I. Binary, sterically accesible interaction previously inferred to be in a complex (number of proteins: MIPS = 145; Pu et al. 2007 = 270)

Type II. Binary, sterically accesible interaction not inferred to be within a complex (MIPS = 97; Pu et al. 2007 = 276)

Type III. Binary, sterically accesible interaction between one protein being in a complex and the other not in the dataset (MIPS = 173 & 317; Pu et al. 2007 = 279 & 350)

Type IV. Binary, sterically accesible interaction between two proteins not in the datasets (MIPS = 727; Pu et al. 2007 = 445)





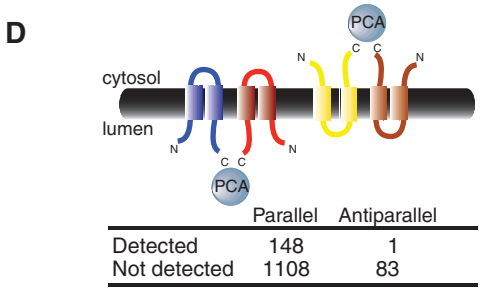


Fig. 3. The DHFR PCA results provide structural and topological insights. PCA fragments have to be in proximity to each other in order to fold into the active structure of the reporter protein. (**A**) PCA PPIs versus protein complexes. Comparison of the PCA network with databases of curated protein complexes (MIPS) and inferred from computational analysis of TAP-MS (*15*) allows classification of four types of PCA interactions: in which both proteins are found within a complex (type 1), are inferred to be in two separate complexes (type 2), one protein is in a complex and the other is not in the network (type 3), or both are absent from the network (type 4) (*15*). Columns of numbers indicate the number of PCA PPIs observed for each data set and each category. (**B**) A thorough DHFR PCA screen of the RNA polymerase II complex [Protein Data Bank (PDB) number 1I3Q] detects predicted interactions among the 10 subunits. (**C**) An interaction is 3.5 times more likely to be detected for a pair of proteins known to interact if the C termini of these proteins are within 82 Å of each other in the case of stable crystallized complexes of yeast-homologous proteins deposited in the PDB. (**D**) Membrane protein topology and PPI detection by PCA. A protein interaction is 12 times more likely to be detected if the C termini are in the same cell compartment.

actions relating information among these processes and CCs that allows their integration into higher-order networks. As we describe below, these interactions reveal specific spatial and topological relationships between known and previously unknown complexes underlying both known and previouly unknown cellular processes.

Global structure and topology of the in vivo interactome. PCA-detected interactions are interpreted differently than purified protein complexes or binary (one-to-one) interactions determined in Y2H screens, allowing us to address how protein complexes and PINs are spatially and topologically organized in living cells. Whether an interaction can be observed depends on the distance between the C termini of two proteins and the length of the polypeptide linker separating bait and prey proteins to the PCA fragments (*12*, *17*) (Fig. 1A). Given that the linkers used in this study were of 10–amino acid residues, for a given protein complex we expected to detect only binary (direct) or near-binary (indirect, C termini within 82 Å, but mediated by one or more other proteins) interactions for protein pairs separated by less than this distance. We first tested this prediction by exhaustively screening all pairwise interactions ($n = 45$ possible pairs) in the well-known RNA polymerase II complex (*10*). We found that we were 5.7 times more likely to detect an interaction if the C-termini were within 82 Å (Fisher's exact test, $P = 0.01$) (Fig. 3B). Interactions that were detected but not predicted could be due to alternative assemblies of this complex in intact cells, to changes in their configuration under different conditions, or to protein dynamics that cannot be interpreted from crystal structures. We then asked whether spatial restraint on observable interactions is reflected in the complete DHFR PCA network. An examination of homologous protein complexes with solved structures showed that we were 3.5 times more likely to detect an interaction between a pair of proteins that have C termini closer than 82 Å than for those with longer distances between C termini (Fisher's exact test, $P < 0.002$) (Fig. 3C). Further, we found that the interacting protein pairs possess domains known to mediate PPI more often than they would be expected to have possessed by chance (7.3% of protein pairs have domains known to mediate PPI versus 0.6%, $P << 0.001$). Thus, the data will be useful for predicting spatial relationships and the bases of molecular recognition among proteins, protein domains, and peptide recognition motifs. Finally, because the C termini of proteins have to be in close proximity and also oriented into the same cellular compartment, PCA provides information about membrane protein topology (Fig. 3D) (*11*, *18*–*20*). Our results demonstrate that the topology of interacting membrane proteins is also reflected in the PIN; specifically, that membrane proteins that colocalize to the same cellular compartment are 12 times more likely to show an interaction if they have a parallel rather than an antiparallel orientation (Fisher's exact test, $P < 0.0005$) (*21*). These PPIs between membrane-associated and membrane-associated and soluble proteins will serve to predict cross-compartment functional relationships, such as interactions of endoplasmic reticulum–associated membrane receptors and cytosolic or nuclear effector proteins.

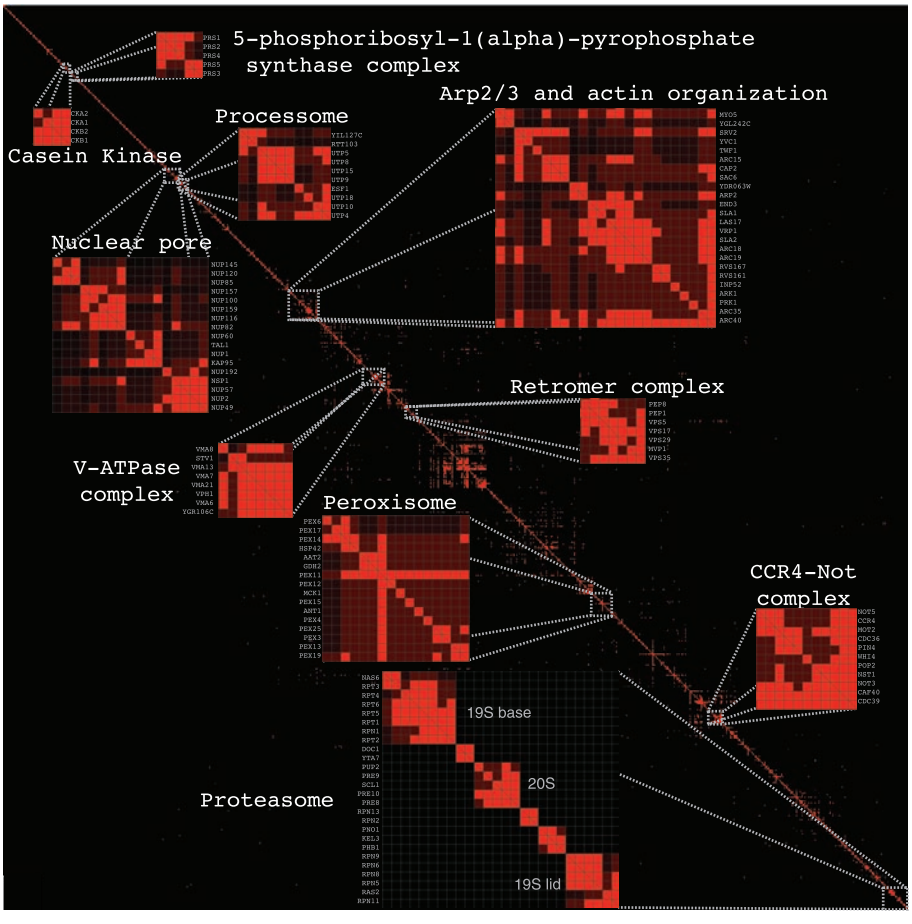


Fig. 4. The DHFR PCA network is modular and interconnected. Clustering of the DHFR PCA network reveals numerous known complexes, within which the substructure represents known subunits. Proteins that have interaction patterns similar to those of other proteins and that interact together are grouped along the diagonal.

Bird's-eye view of the yeast in vivo PIN. The general predictions described above led us to pose specific hypotheses for how protein complexes and networks are organized in living cells. Unsupervised hierarchical clustering of the 2770 DHFR PCA interactions provides an overview of the in vivo PIN (Fig. 4 and file S1). A number of crystallographically or biochemically well-characterized complexes are organized as clusters along the diagonal, confirming that their organization in cells reflects their predicted structures in vitro. Also, substructures of these clusters are consistent with those of previously affinity-purified subcomplexes. For instance, the nuclear pore contains a number of distinct subclusters, three of which clearly correspond to known subfractionated complexes (the Nup84 subcomplex includes Nup85, Nup120, and Nup145, which are in the network, and Nup84 and Seh1p, which are not in the network; a second subcomplex that includes Nup57, Nup49, and Nsp1, which are in the network, and Nic96, which is not in the network; the Nup82 subcomplex includes Nup159, Nup82, and Nsp1, which interacts with Nup166 for its proper localization) (*22*). These subcomplexes also represent groups of proteins that have been hypothesized to form direct contacts in a detailed architectural map of the assembly of the nuclear pore complex (*23*). Our results now suggest that such substructures exist in intact cells. Similarly, the proteasome partitions into subcomplexes that correspond to the composition of characterized fractions and of structures that can be visualized in intact cells (*24*). Complexes described in vitro can therefore accurately reflect those seen in vivo by PCA and as reported by whole-cell electron tomographic studies of protein complexes (*25*).

PPIs between complexes that reflect the cross-compartmental and cross-functional interactions described above (Fig. 2 and fig. S9) can be visualized as off-diagonal interactions on this map (Fig. 4). These represent links among several network modules that have been well described and shown to be central to eukaryotic cell biology. Our map therefore allows us to

identify previously unkown multifunctional PINs and to associate and integrate other proteins to these processes (Fig. 5). First, we showed that, starting from the Arp2-actin organization network, we are able to describe new connectivity among the complex network of interactions that integrates actin filament assembly and patch formation with secretion and cell-wall synthesis and ultimately with membrane abscission and cell separation during cell division. Second, we showed that starting from the retromer complex, we can physically integrate the protein-sorting machinery and trafficking with the synthesis of autophagosomes, links that were previously suggested from genetic and cell-biological observations.

In vivo network at the bud neck. Successful completion of a mitotic cell cycle, and the creation of a viable daughter cell, involves the tight coordination of nuclear events with mechanisms that control cell morphology. The formation of a bud is an example of this temporal and spatial coordination. The bud is the recipient of the segregated material from a mother cell, and ultimately will form the daughter cell. DNA, organelles, proteins, and mRNAs are selectively transported into the bud from the mother and after cytokinesis are enclosed within the daughter cell (*26*). After bud-site selection, recruitment and assembly of proteins that act in polarized growth and bud emergence occur at the incipient bud site. We can construct a network (Fig. 5A) that captures the dynamic assembly and localization of "polarisome" proteins and both known and previously unknown interactions to proteins that provide input signals from the cell-cycle machinery via specific cyclin–cyclin-dependent kinase 1complexes and the Rho signaling proteins that control polarity and cell integrity (*27*). These mechanisms organize and polarize the cytoskeleton and the secretory apparatus at the bud tip and bud neck during cell-cycle progression. Many polarisome proteins that localize to the bud and bud neck (Fig. 5A, blue), or predominantly to the neck (Fig. 5A, orange) are found in the PCA network. Bem1 plays a central role in bud polarization through its ability to build scaffolds, at sites of polarized growth, of an activator (Cdc24) and an effector (Ste20) of the Cdc42 Rho-like guanosine triphosphatase (GTPase) (*28*). Kel1 and Kel2 also act as scaffolds for polarity components at the bud tip and bud neck and were shown to couple to Bem1 and Spa2. Further, the exocyst complex functions in the vectoral transport of vesicles from the Golgi to the bud and promotes plasma membrane expansion, and the Arp2/3 complex, by mediating the assembly of actin patches, promotes membrane recycling through endocytosis (*29*). An extensive network of proteins containing the Arp2/3 actin-assembly complex, its activator Las17, and effectors of actin organization is represented in our network and recapitulates many known protein interactions (Fig. 5A, orange edges), but extends the level of connectivity among components (Fig. 5A, blue edges), especially for Sla2, Las17, and Arc40. However, Arc40 of the Arp2/3 complex is linked via Rvs161 to the GTPase-activating protein (GAP) Gyl1, known to function in actin-patch formation and polarized exocytosis; Gyl1 is also connected to the exocyst through the Sec4 GTPase and to Ynr065C (Fig. 5A), a large protein of previously unknown function and for which localization data are unavailable. In a further extension of this actin patch–assembly complex, we found that Las17 and Myo5, a type I myosin that associates with actin patches, interact with Syp1, a protein implicated in actin cytoskeletal organization. Further, we showed that Syp1 physically associates with multiple proteins, including the bud-neck septins Cdc11 and Shs1 and the cell-surface sensors Mid2 and Wsc2, which activate the cell-integrity pathway through Rom2. Collectively, the interactions among distinct complexes seen by PCA represent a potential regulatory network involved in bud polarization, bud-neck organization, and cytokinesis, a network that captures the dynamic transitions of polarity and exocyst components between the bud tip and the bud neck during the cell cycle.

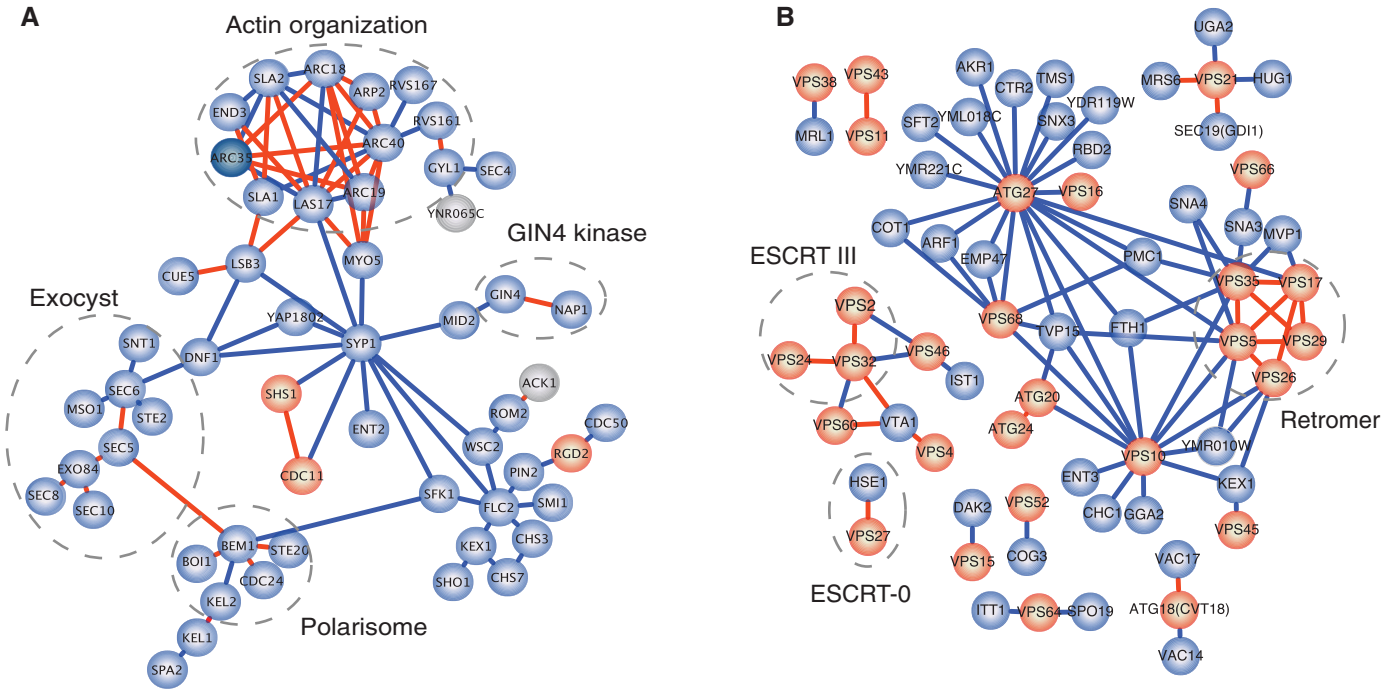


Fig. 5. The yeast DHFR PCA network provides insights into both cell polarity and autophagy. Blue edges denote previously unknown interactions (*10*) and orange edges denote interactions reported at least once in major databases. (**A**) Network at the bud neck. Physical association, detected by PCA between proteins that localize to the bud and bud neck (blue) and proteins that localize predominantly to the bud neck (orange), and proteins lacking localization data (gray), can be used to assemble the structure of a polarity PIN. This PIN shows both known interactions and previously unknown coupling between proteins involved in actin-filament organization and patch and assembly with proteins acting in secretion and cell-wall synthesis. The interactions between protein complexes in the PIN reflect the complex transition of proteins between the bud tip and the bud neck, which function in cell polarity from bud emergence to cytokinesis. (**B**) Autophagy network. Interactions directly connecting proteins involved in autophagy (ATG), vacuolar protein sorting (VPS), and cytoplasm-to-vacuole targeting (CVT) (orange) and other proteins (blue) are shown.

In vivo network of autophagy. Autophagy is the process whereby organelles and the cytosol are engulfed within membrane vesicles for delivery to the lysosome/vacuole for degradation and macromolecule recycling and is involved in development, response to stress, and pathogen resistance. Dysfunction of this conserved eukaryotic process is associated with neurodegenerative conditions, namely Huntington's, Alzheimer's, and Parkinson's diseases, and with cancer (*30*). Proteins involved in autophagy are rich in interactions in the yeast DHFR PCA network (Fig. 5B), including the endosomal sorting complexes required for transport (ESCRTs) ESCRT-0 and ESCRT-III, the retromer complex, and other known interactions (Fig. 5B, dashed gray circles). Vps32, Vps24, and Vps2 are three of the four subunits of the ESCRT-III complex that are responsible for the sorting of transmembrane proteins into the multivesicular body (MVB) pathway. Dysfunction of this complex leads to autophagosome accumulation and neurodegeneration in mammals (*31*). ESCRT-0 is required for sorting ubiquitinated membrane proteins before vacuolar degradation. Vta1 is a member of the MVB pathway and is known to bind to Vps60 and Vps4, regulating the activity of the latter (*31*). Proteins destined for secretion or for delivery to intracellular compartments follow the same route and are sorted in the trans-Golgi network. In yeast, the lysosomal/vacuolar proteins are sorted from other proteins by the carboxypeptidase Y receptor Vps10. These receptors are then returned from the prevacuolar compartments to the trans-Golgi network by the retromer complex. The role of the retromer complex in protein transport is also crucial in metazoa because, for example, it is essential to the formation of important morphogen gradients along body axes (*32*). Some of the interactions we uncovered shed light on this functional relationship between protein sorting and trafficking by the retromer complex and Vps10. For instance, we find that Vps10 shows physical interactions with the retromer complex, which was previously hypothesized on the basis of genetic interactions with Vps35 and Vps26 (*33*, *34*). These observations also suggest a topological relationship between these two proteins that add to our understanding of the structural organization of the retromer complex recently resolved by crystallography (*35*). The interaction between Chc1 (clathrin heavy chain 1 human homolog), the clathrin heavy chain involved in protein transport and endocytosis, and Vps10 was also hypothesized based on genetic data that shows Vps10 is rerouted to the plasma membrane in a *chc1 vps1* mutant instead of its normal travel to the endosome (*36*). We confirm this functional relationship and show that it is mediated through a physical interaction. Most of the interactions we see are previously undescribed (75%) (Fig. 5, blue edges) and represent a substantial advance in describing the autophagy and cytoplasm-to-vacuole targeting pathways. For instance, Atg27 shows a particularly large number of interactions. This protein plays a critical role in the formation of sequestering vesicles, including autophagosomes. It localizes to the Golgi apparatus, the mitochondrion, and the phagophore assembly site. Despite its importance, it showed no interactions in recent large-scale TAP-MS experiments and only one in previous Y2H screens (*3*). Recent work affirms its involvement in both bulk and specific autophagy, and it is hypothesized that Atg27 (along with Atg9, not represented here) labels the membrane source for its transport to and the formation of autophagosomes (*37*). Our results suggest that Atg27 occupies a central role in autophagy because it physically interacts with the retromer complex and with many other vacuolar proteins involved in the sorting of vacuolar hydrolases; further, these results implicate uncharacterized ORFs, such as YML018C, YMR221C, and YDR119W, in this process.

Conclusions. There remain many insights to be drawn beyond the general details, overview, and examples of extended structural and functional networks reported here for the in vivo protein interactome, and other dimensions of the interactome remain to be explored: How dynamic are these interactions? What are the effects of growth conditions on PPI network architecture? The functional and integrative genomic tools developed for this study will enable analysis of protein-interaction dynamics on any scale to uncover mechanisms of biochemical network regulation. A wide variety of PCA reporter enzymes can be used to study temporal and spatial dynamics of protein interactions over a broad range of time scales (from seconds to many hours) and under the influence of natural or artificial perturbations (*9*). Further, the topological requirements of PCA generate a protein-complex topology map at 8-nm resolution that will provide reference data for studying the spatial dynamics of functional protein complexes by immunofluorescence or by monitoring the localization of proteins genetically tagged with fluorescent proteins. Finally, they will also provide reference constraints for determining the architecture of macromolecular assemblies (*23*). The integration of the results from such efforts with those of gene regulation dynamics and protein modifications will lead to a fuller understanding of how complex cellular processes are orchestrated at a molecular and structural level in the living cell.

References and Notes

1. J. Monod, paper presented at the Proceedings of the Eleventh Nobel Symposium, Sodergarn, Lidingo, County of Stockholm, 1968.
2. T. Ito *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 1143 (2000).
3. P. Uetz *et al.*, *Nature* **403**, 623 (2000).
4. T. Ito *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4569 (2001).
5. Y. Ho *et al.*, *Nature* **415**, 180 (2002).
6. A. C. Gavin *et al.*, *Nature* **415**, 141 (2002).
7. A. C. Gavin *et al.*, *Nature* **440**, 631 (2006).
8. N. J. Krogan *et al.*, *Nature* **440**, 637 (2006).
9. S. W. Michnick, P. H. Ear, E. N. Manderson, I. Remy, E. Stefan, *Nat. Rev. Drug Discov.* **6**, 569 (2007).
10. Materials and methods are available as supporting material on *Science* Online.
11. I. Remy, S. W. Michnick, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 5394 (1999).
12. J. N. Pelletier, K. M. Arndt, A. Pluckthun, S. W. Michnick, *Nat. Biotechnol.* **17**, 683 (1999).
13. E. Stefan *et al., Proc. Natl. Acad. Sci. U.S.A.* **104**, 16916 (2007).
14. S. R. Collins *et al.*, *Mol. Cell. Proteomics* **6**, 439 (2007).
15. S. Pu, J. Vlasblom, A. Emili, J. Greenblatt, S. J. Wodak, *Proteomics* **7**, 944 (2007).
16. T. R. Hughes, L. Pena-Castillo, *Genetics* **176**, 7 (2007).
17. I. Remy, I. A. Wilson, S. W. Michnick, *Science* **283**, 990 (1999).
18. M. Diehn, M. B. Eisen, D. Botstein, P. O. Brown, *Nat. Genet.* **25**, 58 (2000).
19. J. P. Miller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12123 (2005).
20. A. A. Zamyatnin Jr. *et al.*, *Plant J.* **46**, 145 (2006).
21. H. Kim, K. Melen, M. Osterberg, G. von Heijne, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 11142 (2006).
22. M. Damelin, P. Silver, *Biophys. J.* **83**, 3626 (2002).
23. F. Alber *et al.*, *Nature* **450**, 683 (2007).
24. O. Medalia *et al.*, *Science* **298**, 1209 (2002).
25. S. Nickell, C. Kofler, A. P. Leis, W. Baumeister, *Nat. Rev. Mol. Cell Biol.* **7**, 225 (2006).
26. V. J. Cid *et al.*, *Microbiology* **148**, 2647 (2002).
27. D. McCusker *et al.*, *Nat. Cell Biol.* **9**, 506 (2007).
28. H. O. Park, E. Bi, *Microbiol. Mol. Biol. Rev.* **71**, 48 (2007).
29. D. Pruyne, A. Legesse-Miller, L. Gao, Y. Dong, A. Bretscher, *Annu. Rev. Cell Dev. Biol.* **20**, 559 (2004).
30. T. Shintani, D. J. Klionsky, *Science* **306**, 990 (2004).
31. J. A. Lee, A. Beigneux, S. T. Ahmad, S. G. Young, F. B. Gao, *Curr. Biol.* **17**, 1561 (2007).
32. D. Y. Coudreuse, G. Roel, M. C. Betist, O. Destree, H. C. Korswagen, *Science* **312**, 921 (2006).
33. S. F. Nothwehr, S. A. Ha, P. Bruinsma, *J. Cell Biol.* **151**, 297 (2000).
34. J. V. Reddy, M. N. Seaman, *Mol. Biol. Cell* **12**, 3242 (2001).
35. A. Hierro *et al.*, *Nature* **449**, 1063 (2007).
36. O. Deloche, R. W. Schekman, *Mol. Biol. Cell* **13**, 4296 (2002).
37. W. L. Yen, J. E. Legakis, U. Nair, D. J. Klionsky, *Mol. Biol. Cell* **18**, 581 (2007).

38. V.M., S.R., M.S., I.S., and Y.M. performed the experiments; S.R. designed the automated robotic procedures; K.T., C.R.L., and S.W.M. analyzed the results; and H.B. helped plan the research and edited the manuscript, which C.R.L., S.W.M., V.M., and J.V. wrote. We thank N. Aubin-Horth and D. Klionsky for comments on the manuscript, J.-F. Paradis for technical and management support, J. Vlasblom for providing scripts and data files, and N. Barkai for providing the gene expression profile compendium.

Supported by grants from Genome Canada, Genome Quebec, and the Canadian Institutes of Health Research (CIHR) (MOP-152556) to S.W.M. S.W.M. holds the Canada Research Chair in Integrative Genomics. C.R.L. is a Natural Sciences and Engineering Research Council of Canada and CIHR Strategic Training Program in Bioinformatics postdoctoral fellow.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1153878/DC1
Material and Methods
Figs. S1 to S11
Tables S1 to S9
File S1
References

7 December 2007; accepted 23 April 2008
Published online 8 May 2008;
10.1126/science.1153878
Include this information when citing this paper.