Polar Mapper: computational tool for integrated visualization of protein interaction networks and mRNA expression data

Joana P. Gonçalves*
Unidade de Sistemas Biológicos, Biocant, Cantanhede, Portugal
and KDBIO Group, INESC-ID, Lisbon, Portugal
and IST, Technical University of Lisbon, Lisbon, Portugal

Mário Grãos*
Unidade de Biologia Celular, Biocant, Cantanhede, Portugal

André X. C. N. Valente†
Unidade de Sistemas Biológicos, Biocant, Cantanhede, Portugal
and Centro de Neurociências e Biologia Celular, Universidade de Coimbra, Portugal

November 5, 2008

Abstract

Polar Mapper is a computational application for exposing the architecture of protein interaction networks. It facilitates the system-level analysis of mRNA expression data in the context of the underlying protein interaction network. Preliminary analysis of a human protein interaction network and comparison of yeast oxidative stress and heat shock gene expression responses are addressed as case studies.

Keywords: protein interaction network; functional genomics; systems biology

Overview

Progress in the reliability and throughput of protein physical interaction detection techniques (both experimental [1–3] and computational wise [4]) is gradually leading to the availability of more comprehensive, higher confidence protein interaction data [5–9]. There is hope that such 'interactome' maps can serve as invaluable tools for biological research, in particular for more integrated system-level studies of biological processes and mechanisms [10]. Notably in this regard, a protein interaction network provides the natural context for interpreting large-scale gene expression data, as the latter can be viewed as the dynamical expression of different parts of the protein interaction network [11, 12]. Now, interactomes can be very large, with rough estimates placing the number of interactions in a human cell on the order of 200 000 [13]. Therefore, for the potential benefits of interactome mapping projects to be realized, proper visualization of interaction data is essential [14]. As an addition to currently available alternatives [15–26], we present in here a software application - Polar Mapper - designed for displaying protein interaction networks in a particularly informative fashion, termed a Polar Map [27]. The software also allows gene expression data to be overlayed on the generated Polar Map for a visually integrated analysis of expression and interaction data. To exemplify the usefulness of Polar Mapper, we applied it to two case studies: i) a preliminary analysis of the collection of

*These authors contributed equally to this work
†Corresponding author

Human protein-protein interaction data obtained via high-throughput Yeast Two-Hybrid (Y2H) assays by Rual *et al.* [1], and ii) a preliminary search for relevant differences in the, a priori very similar [28], expression responses of *S. cerevisiae* to the distinct hydrogen peroxide and heat shock stresses.

Related tools

Several software applications are available for protein interaction network visualization and analysis [15–26]. While typically representing proteins as nodes and interactions as edges, tools such as Pajek [17], BioLayout [23], CNplot [21], PINC [25] or Medusa [20] rely on distinct layout algorithms to produce alternative graphical representations of an interaction network. Spring-embedded, hierarchical, orthogonal, tree and circular are amongst the most widely used kinds of structures [29]. Layouts can be accomplished using a number of distinct approaches, ranging from simulated annealing or gradient-descent minimization of the energy of the representation [23, 24] to hierarchical clustering techniques that group nodes according to their similarity [25]. These techniques are commonly combined with heuristics, which are able to prune the often large set of output arrangements. Moreover, several additional criteria are considered when measuring the quality of the graphical positioning of graph elements: the number of crossing edges, the area occupied by the representation or the distance between adjacent and non-adjacent nodes [29]. Osprey [16], Cytoscape [15], VisANT [18], ProViz [19] and BiologicalNetworks [22] further enable the integration of biological data available at public repositories. Functional annotations can be loaded as node attributes from Gene Ontology [30] by Osprey, Cytoscape, VisANT and BiologicalNetworks. VisANT also supports GenBank [31] and SwissProt [32] annotations. Edges can be complemented with information on pathways and interaction types provided by KEGG [33] database in both VisANT and BiologicalNetworks. The latter further accepts interaction data from the BIND [5] and TransPath [34] databases. Moreover, VisANT enables the integration of homology information and the projection to orthologous genes, based on phylogenetic profiles available at COG [35]. The superposition of gene expression data as additional node information is supported by both Cytoscape and BiologicalNetworks. Additional functionalities for querying, navigating and finding substructures in graphs using either clustering or common algorithms in graphs are also provided by these tools. Interaction network analysis tools are mostly available as standalone applications. BiologicalNetworks, VisANT and Osprey are the exceptions, the first being provided solely as a web-based server and the others supporting both standalone and online execution forms. While the majority of these software tools is implemented in Java thus being compatible with a multitude of platforms, a number of them are restricted to either Windows [17, 24] or UNIX environments [23].

Computational algorithms

Polar Mapper introduces an alternative graphical display of networks, designated a Polar Map, in an application that was developed to be a practical, useful auxiliary tool in biological research projects involving the analysis of protein interaction networks. Additional Polar Mapper software key features include: i) a convenient method for navigating the network based on its modularity analysis; ii) an optional visual superposition of gene expression data upon the interaction network display; iii) the specification of the nodes' sizes as a way to encode further information in the visualization (for instance, the molecular weight of a protein, or the number of members in a protein complex, for cases in which nodes represent protein complexes, rather than individual proteins); iv) the ability to save network information as text and export polar maps as raster (PNG) and vector (SVG and PDF) image files; v) support for maintaining the data and manual annotations associated with a given network in a Polar Mapper session file, enabling users to conveniently have their network analysis work evolve along with their biological research project. Details on how to use the Polar Mapper software in practice are provided in the Polar Mapper Guide, available at the Polar Mapper website [36].

We next provide a description of the Polar Map visualization algorithm [27] integrated in Polar Mapper. An overview of the key steps in the algorithm is shown in Figure 1, while Figure 2 shows an alternative,

more detailed flow chart. Representing the proteins as nodes and the interactions between the proteins as links between those nodes, the question becomes where to place the nodes in the plane, in order to obtain an as meaningful and as visually clear as possible representation of the interaction network. Now, given that the position of each node has two degrees of freedom, this allows the encoding of two distinct types of information in the graph: one, we shall associate with the radial coordinates of the nodes; the other with their angular coordinates.

The radial coordinate is used to introduce a mathematical hierarchical classification for the nodes based on their placement within the network [37]. For this hierarchical classification we choose the betweeness centrality measure [38]. For a node, its betweeness centrality is defined as the total number of shortest paths (between any two other nodes in the network) that pass through it. In keeping with what is visually intuitive, we place higher betweenness centrality nodes closer to the center of the graph and lower betweenness centrality nodes on the periphery of the graph. Due to the long tail of the betweenness centrality values distribution [39] we use a logarithmic scaling, letting the radial coordinate of a node be proportional to $log\big(\max_{BC}/\text{node}_{BC}\big)$, where node_{BC} denotes the betweenness centrality of that node and $\max_{BC}$ the highest betweenness centrality in the network. Noting that proteins range from those that function only within specific well-defined cellular processes, to those that play more global, higher-level functional roles, the inspiration for the above procedure lays in that, perhaps, this biological hierarchy in the role of proteins finds a correspondence in their placement within the mathematical, abstract protein interaction network. The true relevance and form of this parallel between the mathematical network betweenness centrality (or possibly an alternative centrality measure) of proteins and their 'biological hierarchical centrality' is still an open question [40–44]. Regardless, at the least, used in this fashion it is very helpful in visually untangling large protein interaction graphs.

The angular placement of the nodes in the map is going to reflect the modular structure of the mathematical network, in the sense that it contains regions comparatively dense in links. The greedy algorithm of Clauset *et al.* [45] is used to search for a partition of the network into disjoint modules that maximizes the Modularity Score

$$Q = \frac{\text{intra-module links}}{\text{total links}} - \left[\frac{\text{intra-module links}}{\text{total links}}\right]_{\text{random}},$$

where the first term pertains to the network in question as it is, while the second assumes that the links in that network were randomized, subject to every node keeping its original degree. In other words, a high Q score partitioning of the network guarantees that the number of within-module links is maximized with respect to a base random case, represented by the second term in the above formula. Note that the algorithm is not guaranteed to find the partition that yields the Q global maximum [45]. However, the so far significantly modular structure found in protein interaction networks [27, 46–50] assures that in practice the partition found is likely not far off from the optimal one. Combined with the fact that protein interaction networks are large and this algorithm's running time scales almost linearly in the number of nodes for sparse networks (such as protein interaction networks) [45], this makes it a good choice for the purpose at hand. This partitioning of the network is represented visually by allocating each module to a distinct angular region in the graph. That is, the angular coordinates of the nodes are assigned so that all nodes in a given module fall within the same visual conical section. The biological importance of the mathematical partitioning of the network stems from evidence at present supporting that protein modules dense in physical interactions tend to correspond to biological functional modules in the cell [27, 46–49] (see Wang and Zhang [50] for a different view).

Now, the above procedure still leaves the circular ordering of the modules in the graph undetermined. We would like to choose this ordering based solely on the linkage pattern across the modules, placing, to the extent that is possible, closer to each other modules that are in some sense more inter-connected. Formally this is done via the Ring Ordering Algorithm [27], that works as follows. A function that associates an energy E with each potential circular ordering is defined. Given a circular ordering of the modules, let the distance between two modules be the shortest of the two possible distances between them around the circle (i.e., if they are next to each other the distance is 1, if there is a module between them the distance is 2,

etc.). The energy for this circular ordering is then defined as

$$E\,(\text{circular ordering}) = \sum_m \sum_{e_m} d_{e_m}/|e_m|,$$

where, m denotes a module in the network, e_m denotes an edge between module m and another module, d_{e_m} denotes the distance between the modules connected by the e_m edge and $|e_m|$ denotes the total number of edges between module m and other modules. The normalization by $|e_m|$ ensures that every module is given equal weight as far as determining the final arrangement. Now, the lower the energy of an ordering, the better the ordering is considered to be. The search for a low E ordering is done via a greedy procedure. Taking a random circular ordering as a seed, module position permutations are successively checked and performed if they yield a lower E circular ordering. The procedure is also repeated starting from different random seeds, with eventually the net lowest E circular ordering found being the chosen one. Note that the procedure does not guarantee that a global minimum for E is achieved. The angular ordering of the nodes within the angular section assigned to their module is further refined as follows: i) the module is considered as an isolated network (by ignoring links from the module to the rest of the network); ii) the Q-modularity based partitioning is applied to the isolated module producing sub-modules; iii) the Ring Ordering Algorithm is applied to the linkage pattern between these submodules, producing an ordering of the submodules; iv) the angular section of the module is divided amongst the submodules, respecting their ordering from iii); v) within the angular section of a submodule, the angular ordering of its nodes is arbitrary. The motivation for this overall module ordering procedure is again that the density of connections between modules likely correlates with their biological functional closeness, which can be valuably reflected in the graphical display, at least to the extent allowed by the linear circular ordering constraint.

The above algorithm produces a Polar Map for an entire island (isolated graph) in the network (or for the entire network itself, simply by, as a pre-step, assigning separate angular sections to each island). However, it can be useful to visualize Polar Maps of specific regions in a network. A Local Module Polar Map is constructed in the same fashion, upon considering the given module as an isolated network. The modular breakdown into islands, modules and submodules has the additional advantage of providing a structured organization for navigating the network and it is used for that purpose in the Polar Mapper software.

Biological case studies

Human interactome preliminary analysis

The Center for Cancer Systems Biology - Human Interactome 1 (CCSB-HI1) dataset [1] is one of the two first ever collections of human protein-protein interaction data experimentally obtained in a large-scale fashion [1, 2]. In that study, using the yeast two-hybrid assay in a high-throughput format, the products associated with some 8000 human genome open reading frames were systematically pairwise tested for possible physical interactions. This yielded some 2800 binary physical interactions. Note that the large-scale format of the assay is obtained at some cost, for instance, the assay is strictly binary (effects on the interaction of third party proteins or of post-translational modifications are not addressed), and so is its output (interaction detected/not detected, rather than a binding affinity type or other more complex characterization of the interaction). A basic question raised by the above extensive interactome mapping work is how to organize such a large raw dataset: how to grasp its overall structure and how to profitably turn the dataset into a useful aid in specific biological research problems. A more concrete fundamental question is whether indeed some form of functional organization of the cell is present at the level of the interactome, and if it is, whether it is detectable in such a dataset, given the disputed reliability of the yeast two-hybrid technique [13] and the other assay limitations alluded to above.

As our first application of Polar Mapper, we use it as an auxiliary tool in a preliminary exploration of the CCSB-HI1 human interactome dataset. The reader is encouraged to load the associated session file of the human interactome, *HumanInteractome.pm*, and select "Island 1" on Polar Mapper to follow along this

analysis (see supp. file session_intructions for help). Henceforth, module and submodule IDs and names refer to the annotations in this session file. Figure 3 shows the largest connected component of the network. We analyzed the generated modules in order to determine whether they reflected biological functions of the cell. Some modules apparently are not particularly functionally coherent, which may be explained by the fact that the data covers only a very small part of the interactome (on the order of 1% of the existing protein interactions are present in the dataset [1]). False-positive Y2H interactions may also account for some discrepancies [1]. Nevertheless, several modules and submodules clearly show a theme, with the majority of their proteins possessing related functions or sharing common signaling pathways. We could identify modules and/or submodules that are related to Regulation of transcription (mod. 17); House-keeping/biosynthetic pathways (submod. 31); Cell proliferation/death and cancer (mod. 19); Spliceosome/pre-mRNA splicing (submod. 92); Cell division and cancer (submod. 104); Cytoskeleton and protein scaffolding (submod. 52); and Survival signaling (mod. 1).

We now describe an interesting module (mod. 3) we identified whose main theme is membrane-interacting proteins (Figure 4). All its submodules contain membrane-interacting or transmembrane proteins. Three of the submodules should be highlighted in particular, as they are very coherent and can be further sub-categorized as Vesicular transport (I and II) and Secretory pathway/Membrane trafficking (Figure 4). The Vesicular transport I submodule contains well-known SNAREs (soluble N-ethylmaleimide-sensitive factor attachment protein receptors), known for their important role in diverse vesicle-mediated transport events: VAMP4 and VAMP3; syntaxin 4, 5 and 11 ; SNAP23 and 25 [51]. NAPA, also known as α-SNAP, is involved in intra-Golgi transport [51]. SCGN was apparently an outsider, but a recent article described that this protein binds directly to SNAP25 in response to calcium and may be involved in Ca^{2+}-induced exocytotic processes [52]. The SCGN-SNAP25 interaction was not present in the Rual *et al.* [1] Y2H dataset used in this study. On the other hand, a SCGN-SNAP23 interaction was present. It remains to be seen whether this is an artifact of the Y2H screening due to homology of the two SNAPs, or if this interaction is biologically relevant. TAF6L is the only protein unrelated to vesicular transport present in this submodule. The second vesicular transport submodule we highlight contains: RABAC1, involved in vesicle formation from the Golgi complex and that interacts with SNARE complexes [53]; RAB1A, involved in vesicular transport from ER to Golgi [54]; RTN1, shown to bind to several SNAREs [55] and SNX15, involved in endosomal trafficking [56, 57]. It also contains DUSP12, which seems to be unrelated to this submodule's general theme: it is the human ortholog of the *Saccharomyces cerevisiae* Yvh1 protein tyrosine phosphatase [58] and is thought to negatively regulate members of the mitogen-activated protein (MAP) kinase superfamily. The remaining two proteins of the submodule have unknown functions. The third submodule we highlight is composed of proteins involved in secretory pathway/membrane trafficking. Reticulons (RTN1 thru 4; RTN3 and RTN4 are contained in this submodule) are associated with the endoplasmic reticulum and are involved in either neuroendocrine secretion or membrane trafficking in neuroendocrine cells (reviewed in [59]). All members of this family have been shown to interact with and modulate BACE1 (a protease involved in the secretory pathway and a therapeutic target in Alzheimer's Disease). Furthermore, overexpression of any reticulon protein significantly reduces the production of amyloid-beta [60]. RTN3 was described to be involved in membrane trafficking and protein transport between the ER and Golgi [61]. RAB33A is a small GTPase Rab family GTP-binding protein that localizes to dense-core vesicles and may be involved in vesicle transport during exocytosis [62]. LRCH4 is a poorly characterized Leucine-rich protein that contains a carboxyl terminus that may act as a membrane anchor [63], which indicates putative interaction with membranes. PTPN9 is a phosphatase that localises on secretory vesicles [64] and is involved in their fusion control [65]. Finally, COL4A3BP is a kinase involved in non-vesicular ER-to-Golgi transport of ceramide [66], and may be a phosphorylation target of Casein Kinase 1-gamma 2 (CSNK1G2) [67], which is also present in this submodule (their direct physical interaction tested positive in the Rual *et al.* Y2H screen).

For a different example, we now focus on a submodule we found with Polar Mapper, whose interpretation, although less apparent and necessarily more speculative than in the previous case, may lead to interesting findings. In fact, one of the main interests for pursuing large-scale interactome mapping projects is the hope that they can point researchers in a variety of areas to new leads and directions of study. We designated this submodule (submod. 4) as "Crosstalk between toll-like receptors (TLRs) and nuclear receptors". This

submodule is found in a module (mod. 2) containing two other submodules also with proteins that fit in this category (Figure 5). RARs and RXRs are nuclear retinoid receptors that form RAR/RXR heterodimers in response to retinoids (e.g.: retinoic acid), leading to transcription of specific gene networks [68]. RARA, RXRB and RXRG are present in the submodule. SPOP is a poorly characterized protein that is known to bind to and modulate DAXX-mediated transcriptional repression [69]. SPOP was later identified as an adaptor required for the ubiquitination of DAXX by CUL3-based ubiquitin ligase and consequent degradation by the proteasome [70]. DAXX is a multifunctional protein that is involved in a wide variety of processes, such as transcription, cell cycle, and apoptosis [71]. Unfortunately, DAXX was not present in the Y2H dataset [1] used in our study. Members of the nuclear-receptor superfamily repress proinflammatory programs of gene expression. The use of specific agonists for nuclear receptors such as GR (glucocorticoid receptor), LXR (liver X receptors), PPARs (peroxisome proliferator-activated receptors), and to a lesser extent, RARs, were found to modulate both common and distinct subsets of TLR target genes [72]. *DAXX* mRNA expression was over 12-fold upregulated upon stimulation of macrophages using LPS, a well known TLR agonist. In the presence of specific agonists for GR, LXRα/β and PPARγ, the LPS-induced response was inhibited by 48%, 55% and 18%, respectively [72]. Unfortunately this experiment was not done using RARs agonists, since the authors of the study focused on the receptors that modulated the higher number of genes on the initial screening, which does not allow confirmation of whether RARs are involved in the modulation of *DAXX* expression upon LPS stimulus. Nevertheless, the link between retinoic acid receptors and DAXX is still present through RXRs, which may also form heterodimeric pairs with other nuclear receptors besides RARs, such as PPARs and LXR [68], which were shown to modulate *DAXX* expression [72]. DAXX was also shown to negatively modulate the transcriptional activity of Androgen Receptor (another nuclear receptor) [73]. MYD88 (present in the current submodule) is downstream of several TLRs and is involved in innate immunity signaling (namely through the p38 and JNK pathways) [72] and TLR-induced apoptosis, through interaction with FADD [74].

Overall, this preliminary analysis of the CCSB-HI1 dataset served to confirm the presence and relative ease of finding of functional coherent modules in high-throughput interactome data. The latter specific example found and discussed above, also hints at the likely presence of potentially interesting new leads for a variety of biological research areas in these datasets.

Comparative expression analysis of yeast under hydrogen peroxide and heat shock stress

Microarray based high-throughput gene expression profiling assays provide in many regards a similar challenge to high-throughput interactome mapping assays, namely how to handle the associated large quantities of data and how to extract valuable insights from them. Again, as in the interactome field, the extent to which the level of noise in microarray gene expression assays affects the usefulness of the produced data is a point of dispute. It has been shown already that jointly analyzing interaction and expression data can be particularly informative [11, 12, 75]. Polar Mapper has been set up to allow so in the most straightforward possible fashion: by visually superposing the expression data (with the standard green/red color scheme; see Polar Mapper Guide Supp. Mat. file for details on the coloring scheme) over the interactome network. In the example that follows, we use Polar Mapper, the *S cerevisiae* Filtered Yeast Interactome (FYI) [76] and high-throughput expression data from Gasch et al. [28] to probe for differences between the yeast oxidative stress (hydrogen peroxide) and heat shock gene expression responses in yeast. The heat shock and oxidative stress responses in yeast are reported to be very similar [28]. This similarity has been explained as a result of the general environmental stress response (ESR) of yeast [28]. The main difference between the responses to the two stimuli was identified as being associated with a restricted set of genes related to detoxification processes and reductive reactions in the cell [28]. We shall see that the Polar Mapper interactome-based visualization of the expression data can be useful to make apparent other potential differences between these similar expression responses, in particular, differences associated with specific cellular processes, since the modular structure of an interactome appears to reflect such processes.

We start with a Polar Mapper analysis of the observed hydrogen peroxide stress (HP) response by itself.

The reader is encouraged to load the associated session file for the yeast gene expression response to hydrogen peroxide, *Yeast_H2O2_fyi.pm* (showing the expression data superposed on the interactome network data) and select "Island 1" on Polar Mapper at this stage. The module and submodule names and numeric references in the analysis that follows all refer to that annotated Polar Mapper session. Note that most of the names annotating the modules are lifted from the analysis in [27]. Overlaying the mRNA expression data [28] obtained from *S. cerevisiae* under HP stress (0,30 mM during 20 min.) on the FYI yeast interactome using Polar Mapper, it becomes immediately evident that the genes in several modules of the largest connected component of the interactome behave in ensemble, resulting in entire modules having a clear trend towards repression or induction (Figure 6). This pattern is also visible at the submodular level. Some modules present clusters of genes that are up or down-regulated, which are consistent with the submodular grouping [eg.: the Cell cycle control (mod. 23), Signaling (mod. 20), and the RNA processing/translation (mod. 18) modules]. A very robust repression of ribosomal transcripts (see the Large 60S (mod. 6) and Small 40S (mod. 25) ribosomal subunit modules) is apparent. This type of response to stress (including oxidative stress) is well known [28, 77, 78]. Modules composed of genes involved in mRNA related processes are repressed as well, which is evident in the Exosome (mod. 8) and the Spliceosome (mod. 17) modules. Additionally, translation has also been reported to be repressed during stress [28, 79], which can be readily identified in the Translation initiation complex module (mod. 21) and the Translation/Translation initiation submodule (mod. 18, submod. 74). Genes involved in mitosis are also repressed, as seen in the Anaphase promoting complex (APC) submodule (mod. 14, submod. 55); the Chromosome condensation/segregation module (mod. 13); the Cytokinesis/Chromosome segregation module (mod. 22) and submodule 99 of the Cell cycle control module (mod. 23) (analyzed later in greater detail). These results are consistent with reports that HP induces a G2/M arrest in yeast [80]. Conversely, some modules are clearly upregulated. There is induction of genes involved in protein degradation, namely proteasomal genes (see the Proteasomal regulatory complex and the Proteasomal catalytic complex modules, modules 4 and 3, respectively), which is in agreement with other studies using HP as a stressor on yeast [77]. The DNA repair submodule (mod. 14, submod. 57) is also upregulated. It is known that HP and other oxidative stress inducers may generate DNA damage and induce the cellular DNA repair mechanisms [28, 81]. Other modules present upregulated submodules, but overall, the notion that gene repression is predominant under oxidative stress [28] becomes rather evident upon visualization of the overall expression profile superimposed on the interactome (Figure 6). The Cell cycle control module (mod. 23) is an interesting example of one module that does not show a clear trend towards up or down-regulation. However, by zooming in at the module level, it becomes evident that it contains submodules that are induced, while others are repressed. One of the submodules, named the G1/S submodule (submod. 98) contains mainly proteins that are involved in the G1 and S phases of the cell cycle. This submodule is upregulated, which is in agreement with results showing that yeast cells are able to progress through G1/S upon HP challenge [80]. Interestingly, this study also showed that the S phase duration was slightly prolonged compared to untreated cells. This fact is consistent with the observation that, despite many G1/S transcripts being upregulated, *CDC6* is downregulated. This gene is essential for DNA replication initiation during S phase [82]. Several G2/M-related proteins can be found in submodule 99 in which, with the exception of *CAK1*, all the G2/M related genes are downregulated. This reinforces the strong effect of HP on this phase of the cell cycle.

We now compare the heat shock (HS) (25 to 37°C during 20 min, data from [28]) and hydrogen peroxide (HP) stress responses of yeast. A key technique we use here is to visualize the hydrogen peroxide (HP) stress expression response *relative* to the heat shock (HS) stress expression response. This is done by loading into Polar Mapper the difference between the log expression data for the HP and HS responses (Figure 7; Polar Mapper session *Yeast_H2O2-Heat_Shock_fyi.pm, Island 1*). For the most part, the trend towards induction or repression in any given interactome module is the same under both HP and HS stresses. But by looking at the above difference, the modules that are *more intensely* induced or *more intensely* repressed under HP than under HS become evident, since they retain their net original color associated with up (red) or down-regulation (green) respectively, observed in the HP-only image (Figure 6). Conversely, when the color trend is reversed between Figure 6 and Figure 7, it must be that the induction/repression is more intense in HS than in HP. It becomes therefore simple to visually identify modules and submodules that may be more

affected or regulated by each specific stress.

From Figures 6 and 7 it becomes evident that some modules are more strongly upregulated or downregulated by each particular stimulus (see Table 1, HP vs. HS). During the HP challenge, the repression of the G2 and mitosis related modules is stronger than in HS stress: see the Chromosome condensation/segregation module (mod. 13), the APC submodule (submod. 55) and the Cytokinesis/chromosome segregation module (mod. 22). The Cell cycle control module (mod. 23) provides us with further information: the G2/M inducers are more repressed or less upregulated in HP compared to HS (Table 2). This suggests that HP has an effect on the cell cycle, namely on the G2/M phase of the cell cycle. This is in agreement with published data, reporting the existence of a G2/M block after HP stress [80], but not during HS [83]. In fact, this heat shock study [83] reports that HS induces a transient G1/S arrest. This is supported by data showing that several key G1/S genes are differentially expressed between HP and HS: the expression of the G1/S inhibitor *SIC1* remains unchanged in HP and is upregulated in HS; conversely, the G1/S inducers *CLB5* and *CLB6* (critical for DNA replication) remain unchanged in HP and repressed in HS; *SWI4* (involved in G1/S progression) is also less strongly upregulated in HS then in HP. Although there are some exceptions and also some incomplete data, in general, the G1/S gene expression comparison on this module fits the experimental data (see Table 2). Interestingly, the DNA repair submodule (submod. 57 and Table 1) is more strongly induced in HP than in HS, which is also in line with the general notion that direct DNA damage is extremely important during oxidative stress [81, 84, 85]. Overall, this analysis suggests a stronger control of protein and RNA synthesis and degradation and G1/S progression upon heat shock and mitotic control and DNA repair upon HP stress (Table 1).

Due to the small number of chaperones present in the FYI dataset we analyzed, a relevant matter we did not explore is the collective change induced by stress in chaperones and their low-affinity, but fundamental, transient interactions [86]. For instance, similarly based on a combined gene expression and protein interaction data analysis, it has been hypothesized that in yeast cellular stress leads chaperones to become more central in the interactome [75]. A next step in validating and refining some of the ideas related to the dynamic nature of the interactome will likely require the actual experimental testing of protein-protein interactions in cells under different conditions. It has also been shown recently that in eukaryotic cells, in contrast to bacteria, there are two distinct chaperone networks, one being involved in de novo protein folding (coupled to translation), and the other in rescue of stress-denatured proteins [87]. In this regard, from the 14 chaperones identified as translation-coupled by Albanèse *et al.* [87], 7 out of the 8 present in the FYI dataset we analyzed, belonged to the same interactome module (mod. 23, submod. 100; the exception being 1 protein, present in mod. 23, submod. 101). Conversely, out of the 20 chaperones categorized as stress-coupled by Albanèse *et al.* [87], the 4 presented in the FYI dataset were all grouped together by PolarMapper in a distinct module from the translation-coupled one (mod. 10, submod. 37). The placement in the interactome of these few chaperones present in the FYI dataset, seems therefore to be consistency with the two separate functional chaperone classes identified in the study of Albanèse *et al.* [87].

Overall, our case study showed how this graphical platform, combining expression and interaction data, can aid in a first-pass analysis and organization of expression data. Note in particular how it allowed a faster identification of relevant cellular processes, giving hints on biological processes that should be subjected to a more detailed study. It also serves to note that, in spite of the respective assay reliability limitations, present day high-throughput interaction and expression data can be already valuable resources in biological research.

Summary

This article introduces Polar Mapper, a computational application centered around the Polar Map visualization of protein interaction networks. It is meant to be a practical, ready-to-use auxiliary tool in biological research work that involves the analysis of protein interaction networks. A second objective of the project is to make available to the scientific community a reusable implementation of the Polar Map algorithm in order to contribute to the ongoing development of improved biological network visualization tools. To this

end the source code and documentation are open and freely available to the community. Finally, although visualization of protein interaction networks was the primary motivation for this application, we note that it may likely be profitably employed in the display of many other kinds of binary interaction data networks.

Availability

Polar Mapper is implemented in Java and may be used within several platforms and environments, as long as a Java Virtual Machine installation is provided. Binaries and source code, as well as documentation are available both as electronic supplementary materials at the Journal of the Royal Society Interface website and at the Polar Mapper website [36].

Electronic Supplementary Material Files

Session_Instructions.pdf — Instructions for using Polar Mapper with the article case-study sessions

Quick instructions on how to work with the Polar Mapper sessions.

HumanInteractome.pm — Human interactome session

Polar Mapper session containing Human interactome based on Y2H dataset of Rual et al. [1].

Yeast_H2O2_fyi.pm — Yeast hydrogen peroxide oxidative stress session

Polar Mapper session containing yeast interactome based on the FYI dataset of Han et al. [76] together with the hydrogen peroxide oxidative stress expression response data of Gasch et al. [28].

Yeast_H2O2-Heat_Shock_fyi.pm — Yeast hydrogen peroxide oxidative stress expression response *relative* to heat shock expression response session

Polar Mapper session containing yeast interactome based on the FYI dataset of Han et al. [76] together with data showing the difference between the log expression for the hydrogen peroxide oxidative stress response and the log expression for the heat shock response (expression data from Gasch et al. [28]).

PolarMapper_Guide.pdf — Guide to the Polar Mapper software

Polar Mapper Guide providing information on software pre-requisites, installation instructions and usage scenarios.

Acknowledgements

The authors would like to thank the invaluable and professional administrative support at Biocant provided by: Margarida Catarino, Sandra Fernandes, Cândida Baptista, Sandra Tomé, Vera Ganilho and Mariana Brandão (administrative), and António Sampaio and Paulo Alves (IT). The work of JPG was partially supported by Biocant funding and by FCT grant SFRH/BD/36586/2007.

References

[1] Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, *et al*: **Towards a proteome-scale map of the human protein-protein interaction network**. *Nature* 2005, **437**(7062):1173–1178.

[2] Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, S K, *et al*: **A human protein-protein interaction network: A resource for annotating the proteome**. *Cell* 2005, **122**(6):957–968.

[3] Gingras AC, Gstaiger M, Raught B, Aebersold R: **Analysis of protein complexes using mass spectrometry**. *Nature Rev Mol Cell Biol* 2007, **8**:645–654.

[4] Valencia A, Pazos F: **Computational methods for the prediction of protein interactions**. *Curr. Opin. Struct. Biol.* 2002, **12**:368–373.

[5] Bader GD, Betel D, Hogue CWV: **BIND: the Biomolecular Interaction Network Database**. *Nucleic Acids Research* 2003, **31**:248–250.

[6] Xenarios I, Fernandez E, Salwinski L, Duan XJ, Thompson MJ, Marcotte EM, Eisenberg D: **DIP: The Database of Interacting Proteins: 2004 update**. *Nucleic Acids Research* 2004, **32 (database issue)**:D449–D451.

[7] Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M: **BioGRID: a general repository for interaction datasets**. *Nucleic Acids Research* 2006, **34 (database issue)**:D535–D539.

[8] Wu X, Zhu L, Guo J, Fu C, Zhou H, Dong D, Li Z, Zhang DY, Lin K: **SPIDer: Saccharomyces protein-protein interaction database**. *BMC Bioinformatics* 2006, **7**(Suppl. 5):S16.

[9] Mishra GR, Suresh M, Kumaran K, Kannabiran N, Suresh S, Bala P, Shivakumar K, Anuradha N, Reddy R, M RT, *et al*: **Human protein reference database - 2006 update**. *Nucleic Acids Research* 2006, **34 (database issue)**:D411–D414.

[10] Uetz P, Finley Jr RL: **From protein networks to biological systems**. *Febs Letters* 2005, **579**:1821–1827.

[11] de Lichtenberg U, Jensen LJ, Brunak S, Bork P: **Dynamic Complex Formation During the Yeast Cell Cycle**. *Science* 2005, **307**(5710):724 – 727.

[12] Jansen R, Greenbaum D, Gerstein M: **Relating whole-genome expression data with protein-protein interactions**. *GenomeResearch* 2002, **12**:37–46.

[13] Hart GT, Ramani AK, Marcotte EM: **How complete are current yeast and human protein-interaction networks?** *Genome Biology* 2006, **7**:120.

[14] Hu Z, Mellor J, Wu J, Kanehisa M, Stuart JM, DeLise C: **Towards zoomable multidimensional maps of the cell**. *Nature Biotechnology* 2007, **25**(5):547–554.

[15] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: **Cytoscape: A software environment for integrated models of biomolecular interaction networks**. *Genome Research* 2003, **13**:2498–2504.

[16] Breitkreutz BJ, Stark C, Tyers M: **Osprey: a network visualization system**. *Genome Biology* 2003, **4**:R2.

[17] Batagelj V, Mrvar A: **Pajek - Program for large network analysis**. *Connections* 1998, **21**:47–57.

[18] Hu Z, Mellor J, Wu J, DeLisi C: **VisANT: an online visualization and analysis too for biological interaction data**. *BMC Bioinformatics* 2004, **5**:17.

[19] Iragne F, M N, Mathieu B, D A, D S: **ProViz: protein interaction visualization and exploration**. *Bioinformatics* 2005, **21**(2):272–274.

[20] Hooper SD, Bork P: **Medusa: a simple tool for interaction graph analysis**. *Bioinformatics* 2005, **21**(24):4432–4433.

[21] Batada NN: **CNplot: visualizing pre-clustered networks**. *Bioinformatics* 2004, **20**(9):1455–1456.

[22] Baitaluk M, Sedova M, Ray A, Gupta A: **BiologicalNetworks: visualization and analysis tool for systems biology**. *Nucleic Acids Research* 2006, **34**:W466–W471 Sp. Iss. SI.

[23] Enright AJ, Ouzounis CA: **BioLayout - an automatic graph layout algorithm for similarity visualization**. *Bioinformatics* 2001, **17**(9):853–854.

[24] Li W, Kurata H: **A grid layout algorithm for automatic drawing of biochemical networks**. *Bioinformatics* 2005, **21**(9):2036–2042.

[25] Lu H, Zhu X, Liu H, Skogerbø G, Zhang J, Zhang Y, Cai L, Zhao Y, Sun S, J X, *et al*: **The interactome as a tree - an attempt to visualize the protein-protein interaction network in yeast**. *Nucleic Acids Research* 2004, **32**(16):4804–4811.

[26] **PIMWalker** [http://pimr.hybrigenics.com/PIMRider/PIMRider-Categorie/PIMWALKER.html].

[27] Valente AXCN, Cusick M: **Yeast Protein Interactome topology provides framework for coordinated-functionality**. *Nucleic Acids Research* 2006, **34**(9):2812–2819.

[28] Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO: **Genomic expression programs in the response of yeast cells to environmental changes**. *Mol Biol Cell* 2000, **11**:4241–4257.

[29] Tollis I, Di Battista G, Eades P, Tamassia R: *Graph Drawing: Algorithms for the Visualization of Graphs*. Prentice-Hall 1998.

[30] Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G: **Gene Ontology: tool for the unification of biology**. *Nature Genetics* 2000, **25**:25–29.

[31] Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL: **GenBank**. *Nucleic Acids Research* 2007, **35**:D21–D25.

[32] Wu C, Apweiler R, Bairoch A, Natale D, Barker W, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin M, Mazumder R, O'Donovan C, Redaschi N, Suzek B: **The Universal Protein Resource (UniProt): an expanding universe or protein information**. *Nucleic Acids Research* 2006, **34**:D187–D191.

[33] Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M: **KEGG: Kyoto Encyclopedia of Genes and Genomes**. *Nucleic Acids Research* 1999, **27**:29–34.

[34] Krull M, Pistor S, Voss N, Kel A, Reuter I, Kronenberg D, Michael H, Schwarzer K, Potapov A, Choi C, Kel-Margoulis O, Wingender E: **TRANSPATH ®: an information resource for storing and visualizing signaling pathways and their pathological aberrations**. *Nucleic Acids Research* 2006, **34**:D546–D551.

[35] Tatusov R, Natale D, Garkavtsev I, Tatusova T, Shankavaram U, Rao B, Kiryutin B, Galperin M, Fedorova N, Koonin E: **The COG database: new developments in phylogenetic classification of proteins from complete genomes**. *Nucleic Acids Research* 2001, **29**:22–28.

[36] **Polar Mapper website** [http://kdbio.inesc-id.pt/software/polarmapper].

[37] Brandes U: **Communicating centrality in policy network drawings**. *IEEE Trans. Vis. Comput. Graph* 2003, **9**:241–253.

[38] Freeman L: **A set of measures of centrality based upon betweenness**. *Sociometry* 1977, **40**:35–41.

[39] Goh KI, Kahng B, Kim D: **Universal behavior of load distribution in scale-free networks**. *Physical Review Letters* 2001, **87**:Art. No. 278701.

[40] Joy MP, Brock A, Ingber DE, Huang S: **High betweenness proteins in the yeast interaction network**. *J Biomed Biotechnol* 2005, **2**:96–103.

[41] Hahn MW, Kern AD: **Comparative genomics of centrality and essentiality in three eukaryotic protein interaction networks**. *Mol Biol Evol* 2005, **22**:803–806.

[42] Coulomb S, Bauer M, Bernard M, Marsolier-Kergoat MC: **Gene essentiality and the topology of protein interaction networks**. *Proc of the Royal Society B - Biological Sciences* 2005, **272**(1573):1721–1725.

[43] Estrada E: **Virtual identification of essential proteins within the protein interaction network of yeast**. *Proteomics* 2006, **6**:35–40.

[44] Junker BH, Koschützki D, Schreiber F: **Exploration of biological network centralities with CentiBiN**. *BMC Bioinformatics* 2006, **7**:219.

[45] Clauset A, Newman MEJ, More C: **Finding community structure in very large networks**. *Physical Review E* 2004, **70**:Art. No. 066111.

[46] Spirin V, A M: **Protein complexes and functional modules in molecular networks**. *Proc Natl Acad Sci USA* 2003, **21**:12123–12128.

[47] Pereira-Leal JB, Enright JA, A OC: **Detection of functional modules from protein interaction networks**. *Proteins* 2004, **54**:49–57.

[48] Brun C, Herrnann C, Guénoche A: **Clustering proteins from interaction networks for the prediction of cellular functions**. *BMC Bioinformatics* 2004, **5**:Art. No. 95.

[49] Nabieva E, Jim K, Agarwal A, Chazelle B, Singh M: **Whole-proteome prediction of protein function via graph-theoretic analysis of interaction maps**. *Bioinformatics* 2005, **21**:I302–I310.

[50] Wang Z, Zhang J: **In Search of the Biological Significance of Modular Structures in Protein Networks**. *PLoS Computational Biology* 2007, **3**(6):e107.

[51] Hong W: **SNAREs and traffic**. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 2005, **1744**:120–144.

[52] Rogstam A, Linse S, Lindqvist A, James P, Wagner L, Berggard T: **Binding of calcium ions and SNAP-25 to the hexa EF-hand protein secretagogin**. *Biochem J* 2007, **401**:353–363.

[53] Gougeon PY, Prosser DC, Da-Silva LF, Ngsee JK: **Disruption of Golgi Morphology and Trafficking in Cells Expressing Mutant Prenylated Rab Acceptor-1**. *J. Biol. Chem.* 2002, **277**:36408–36414.

[54] Tisdale EJ, Bourne JR, Khosravi-Far R, Der CJ, Balch WE: **GTP-binding mutants of rab1 and rab2 are potent inhibitors of vesicular transport from the endoplasmic reticulum to the Golgi complex**. *J Cell Biol* 1992, **119**:749–761.

[55] Steiner P, Kulangara K, Sarria JCF, Glauser L, Regazzi R, Hirling H: **Reticulon 1-C/neuroendocrine-specific protein-C interacts with SNARE proteins**. *Journal of Neurochemistry* 2004, **89**:569–580.

[56] Phillips SA, Barr VA, Haft DH, Taylor SI, , Haft CR: **Identification and Characterization of SNX15, a Novel Sorting Nexin Involved in Protein Trafficking**. *J Biol Chem* 2001, **276**:5074–5084.

[57] Barr VA, Phillips SA, Taylor SI, Haft CR: **Overexpression of a Novel Sorting Nexin, SNX15, Affects Endosome Morphology and Protein Trafficking**. *Traffic* 2000, **1**:904–916.

[58] Muda M, Manning ER, Orth K, Dixon JE: **Identification of the Human YVH1 Protein-tyrosine Phosphatase Orthologue Reveals a Novel Zinc Binding Domain Essential for in Vivo Function**. *J Biol Chem* 1999, **274**:23991–23995.

[59] Oertle T, Schwab ME: **Nogo and its paRTNers**. *Trends Cell Biol* 2003, **13**:187–194.

[60] He W, Lu Y, Qahwash I, Hu XY, Chang A, Yan R: **Reticulon family members modulate BACE1 activity and amyloid-beta peptide generation**. *Nat Med* 2004, **10**:959–965.

[61] Wakana Y, Koyama S, Nakajima K, Hatsuzawa K, Nagahama M, Tani K, Hauri HP, Melancon P, Tagaya M: **Reticulon 3 is involved in membrane trafficking between the endoplasmic reticulum and Golgi**. *Biochem Biophys Res Commun* 2005, **334**:1198–1205.

[62] Tsuboi T, Fukuda M: **Rab3A and Rab27A cooperatively regulate the docking step of dense-core vesicle exocytosis in PC12 cells**. *J Cell Sci* 2006, **119**:2196–2203.

[63] Glockner G, Scherer S, Schattevoy R, Boright A, Weber J, Tsui LC, Rosenthal A: **Large-Scale Sequencing of Two Regions in Human Chromosome 7q22: Analysis of 650 kb of Genomic Sequence around the EPO and CUTL1 Loci Reveals 17 Genes**. *Genome Res.* 1998, **8**:1060–1073.

[64] Saito K, Williams S, Bulankina A, Honing S, , Mustelin T: **Association of protein-tyrosine phosphatase MEG2 via its Sec14p homology domain with vesicle-trafficking proteins**. *J Biol Chem* 2007, **282**:15170–15178.

[65] Huynh H, Bottini N, Williams S, Cherepanov V, Musumeci L, Saito K, Bruckner S, Vachon E, Wang X, Kruger J, *et al*: **Control of vesicle fusion by a tyrosine phosphatase**. *Nat Cell Biol* 2004, **6**:831–839.

[66] Hanada K, Kumagai K, Yasuda S, Miura Y, Kawano M, Fukasawa M, Nishijima M: **Molecular machinery for non-vesicular trafficking of ceramide**. *Nature* 2003, **426**:803–809.

[67] Kumagai K, Kawano M, Shinkai-Ouchi F, Nishijima M, Hanada K: **Interorganelle trafficking of ceramide is regulated by phosphorylation-dependent cooperativity between the PH and START domains of CERT**. *J Biol Chem* 2007, **282**:17758–17766.

[68] Bastien J, Rochette-Egly C: **Nuclear retinoid receptors and the transcription of retinoid-target genes**. *Gene* 2004, :1–16.

[69] La M, Kim K, Park J, Won J, Lee JH, Fu YM, Meadows GG, Joe CO: **Daxx-mediated transcriptional repression of MMP1 gene is reversed by SPOP**. *Biochem Biophys Res Commun* 2004, **320**:760–765.

[70] Kwon JE, La M, Oh KH, Oh YM, Kim GR, Seol JH, Baek SH, Chiba T, Tanaka K, Bang OS, Joe CO, Chung CH: **BTB domain-containing speckle-type POZ protein (SPOP) serves as an adaptor of Daxx for ubiquitination by Cul3-based ubiquitin ligase**. *J Biol Chem* 2006, **281**:12664–12672.

[71] Salomoni P, Khelifi AF: **Daxx: death or survival protein?** *Trends Cell Biol* 2006, **16**:97–104.

[72] Ogawa S, Lozach J, Benner C, Pascual G, Tangirala RK, Westin S, Hoffmann A, Subramaniam S, David M, Rosenfeld MG, Glass CK: **Molecular determinants of crosstalk between nuclear receptors and toll-like receptors**. *Cell* 2005, **122**:707–721.

[73] Lin DY, Fang HI, Ma AH, Huang YS, Pu YS, Jenster G, Kung HJ, Shih HM: **Negative modulation of androgen receptor transcriptional activity by Daxx**. *Mol Cell Biol* 2004, **24**:10529–10541.

[74] Aliprantis AO, Yang RB, Weiss DS, Godowski P, Zychlinsky A: **The apoptotic signaling pathway activated by Toll-like receptor-2**. *EMBO J* 2000, **19**:3325–3336.

[75] Palotai R, Szalay MS, Csermely P: **Chaperones as integrators of cellular networks: changes of cellular integrity in stress and diseases**. *IUBMB Life* 2008, **60**:10–15.

[76] Han JD, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJ, Cusick ME, Roth FP, Vidal M: **Evidence for dynamically organized modularity in the yeast protein-protein interaction network**. *Nature* 2004, **430**:88–93.

[77] Marques M, Mojzita D, Amorim MA, Almeida T, Hohmann S, Moradas-Ferreira P, Costa V: **The Pep4p vacuolar proteinase contributes to the turnover of oxidized proteins but PEP4 overexpression is not sufficient to increase chronological lifespan in Saccharomyces cerevisiae**. *Microbiology* 2006, **152**:3595–3605.

[78] Warner JR: **The economics of ribosome biosynthesis in yeast**. *Trends Biochem Sci* 1999, **24**:437–440.

[79] Shenton D, Smirnova JB, Selley JN, Carroll K, Hubbard SJ, Pavitt GD, Ashe MP, Grant CM: **Global translational responses to oxidative stress impact upon multiple levels of protein synthesis**. *J Biol Chem* 2006, **281**:29011–29021.

[80] Shapira M, Segal E, Botstein D: **Disruption of yeast forkhead-associated cell cycle transcription by oxidative stress**. *Mol Biol Cell* 2004, **15**:5659–5669.

[81] Ikner A, Shiozaki K: **Yeast signaling pathways in the oxidative stress response**. *Mutat. Res* 2005, **569**:13–27.

[82] Speck C, Chen Z, Li H, Stillman B: **ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA**. *Nat Struct Mol Biol* 2005, **12**:965–971.

[83] Li X, Cai M: **Recovery of the yeast cell cycle from heat shock-induced G(1) arrest involves a positive regulation of G(1) cyclin expression by the S phase cyclin Clb5**. *J Biol Chem* 1999, **274**:24220–24231.

[84] Wang D, Kreutzer DA, Essigmann JM: **Mutagenicity and repair of oxidative DNA damage: insights from studies using defined lesions**. *Mutat. Res* 1998, **400**:99–115.

[85] Pan X, Ye P, Yuan DS, Wang X, Bader JS, Boeke JD: **A DNA integrity network in the yeast Saccharomyces cerevisiae**. *Cell* 2006, **124**:1069–1081.

[86] Korcsmáros T, Kovács IA, Szalay MS, Csermely P: **Molecular chaperones: The modular evolution of cellular networks**. *J Biosci* 2007, **32**(3):441–446.

[87] Albanèse V, Yam AYW, Baughman J, Parnot C, Frydman J: **Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells**. *Cell* 2006, **124**:75–88.

[88] Newman MEJ, Girvan M: **Finding and evaluating community structure in networks**. *Physical Review E* 2004, **69**:Art. No. 026113.

	Hydrogen peroxide	Heat shock
Stronger downregulation	Chromosome condensation/segregation Cytokinesis/Chromosome segregation Anaphase promoting complex (submod.55)*	Large 60S ribosomal subunit Small 40S ribosomal subunit Translation initiation complex Exosome complex Nucleoplasmic RNA/protein transport Core RNA polymerase
Stronger upregulation	DNA repair submodule (submod. 57)** G1/S submodule (submod. 98)	Protein folding submodule (submod. 100) Proteasomal catalytic complex

* The APC submodule does not show a clear trend towards downregulation during heat shock

** The DNA repair submodule in fact shows a downregulation trend during heat shock

Table 1: Summary of the main modules and submodules identified using Polar Mapper as being more strongly upregulated or downregulated under the hydrogen peroxide or heat shock stimulus in yeast. See main text for analysis.

	HP	HS
G1/S inhibitors		
SIC1	unchanged	+
G1/S inducers		
CLB5	unchanged	-
CLB6	unchanged	-
SWI4 *	+ +	+
SWI6 **	- -	-
MBP1	-	N/A
CLN1	-	- -
CLN2 **	- -	-
G2/M inducers		
CKS1	N/A	-
CLB1	- -	-
CLB2	-	N/A
CLB3	- -	-
CLB4	-	-
CAK1	+	+ +

* SWI4 is involved in DNA synthesis and also in DNA repair

** Genes whose expression is opposite to the trend of most genes in the submodule

CDC28 is also present in this submodule, but it is involved in both G1/S and G2/M progression.

Table 2: Comparative expression of cell cycle genes present in the Cell Cycle control module upon hydrogen peroxide treatment and upon heat shock in yeast. See main text for analysis.

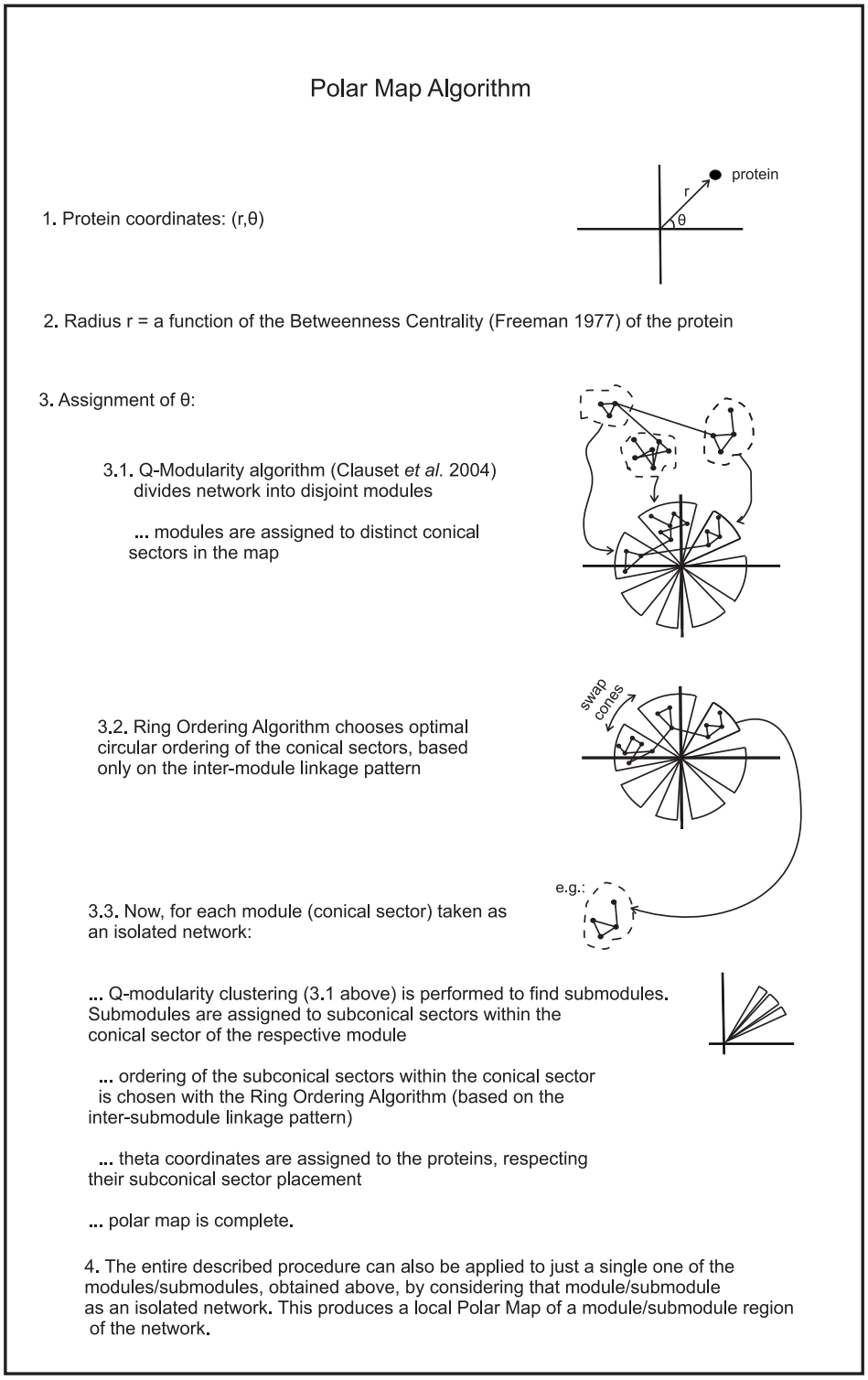
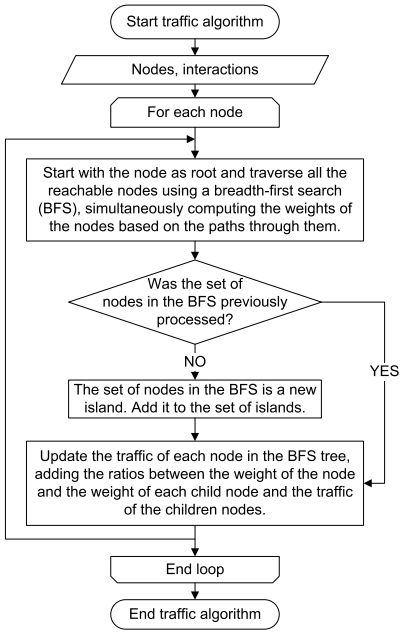
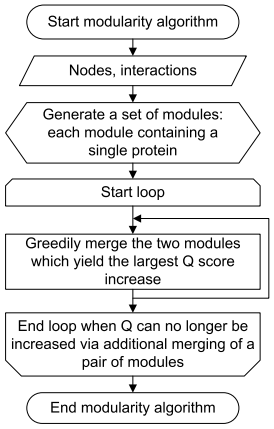


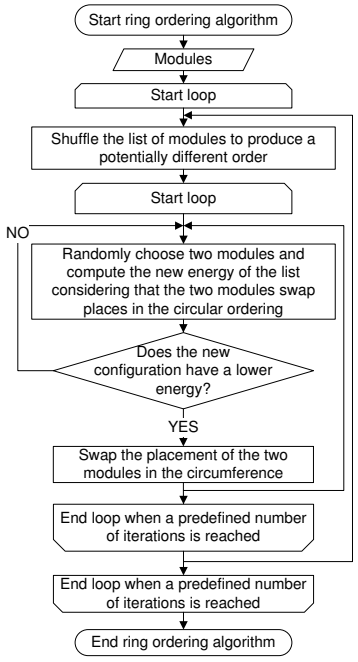
Figure 1: *The Polar Map algorithm.* Overview of the key steps in constructing a Polar Map display of a network.



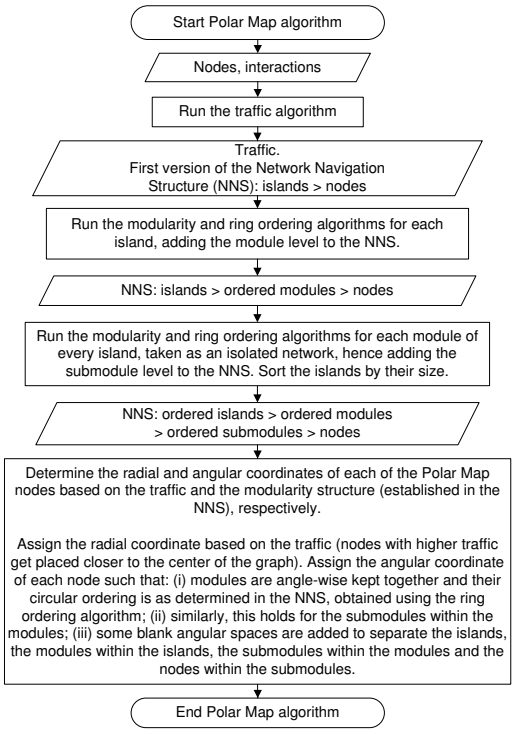
(a) Traffic



(b) Modularity



(c) Ring Ordering



(d) Polar Map

Figure 2: *Detailed flow diagrams of the algorithms used to generate a Polar Map.* (a) Traffic computes the betweeness centrality of each node [38,88] and divides disconnected groups of nodes in islands; (b) modularity generates modules via the Q-modularity algorithm [45]; (c) The ring ordering procedure [27] finds a good ordering for the modules; and, finally, (d) the Polar Map algorithm calculates the polar coordinates of each node based on the results from (a), (b) and (c).

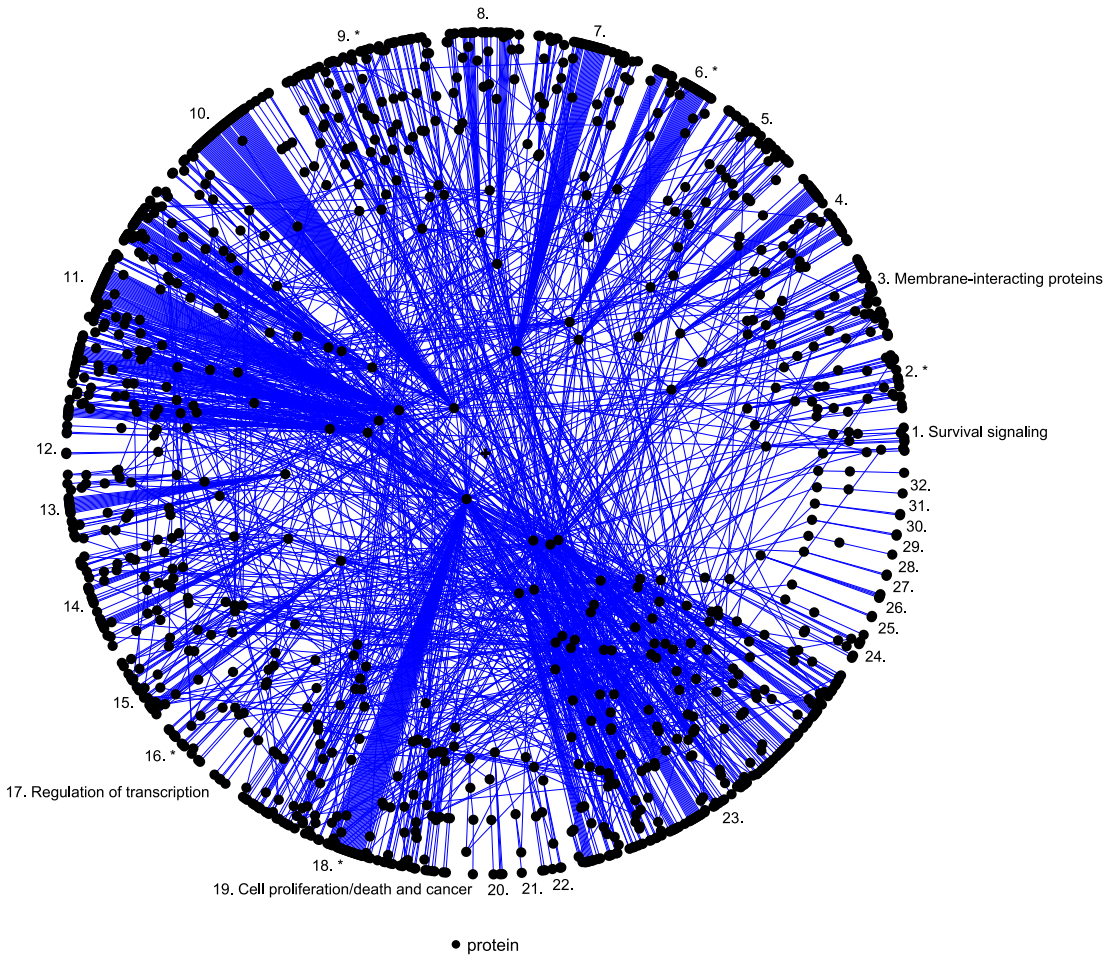


Figure 3: *The Human Interactome.* Polar Map showing the largest connected component of the Human Interactome based on the Rual *et al.* Yeast Two-Hybrid high-throughput protein-protein interaction dataset [1] (1307 proteins, 2441 interactions). Modules are numbered from 1 to 32. Some of them were manually annotated (using Polar Mapper), reflecting the biological function of the proteins that constitute them. Modules marked with an asterisk (*) contain annotated submodules.

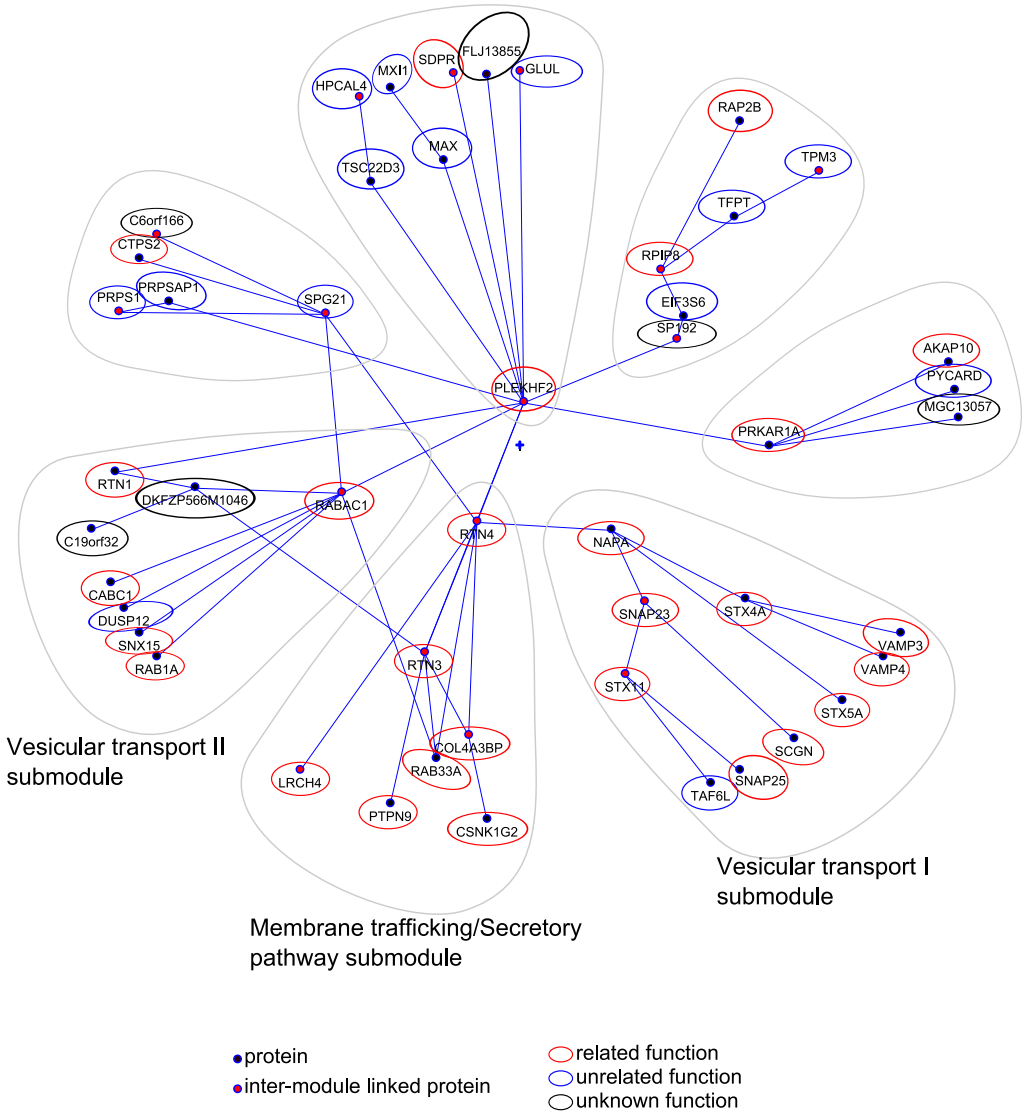


Figure 4: *Membrane-interacting proteins module within the Human Interactome.* Local Polar Map of module 3 in Figure 3. Its central theme appears to be membrane-interacting proteins. Note how the submodules (encircled by grey lines) denser in these proteins are found closer together in the circular ordering of the map. See main text for a biological analysis of this module. To compose this figure, the Polar Map was exported into a SVG file using Polar Mapper and the graphical highlight around the sumodules and relatedness of each protein concerning the module's theme were edited using an external image processing software (Inkscape).

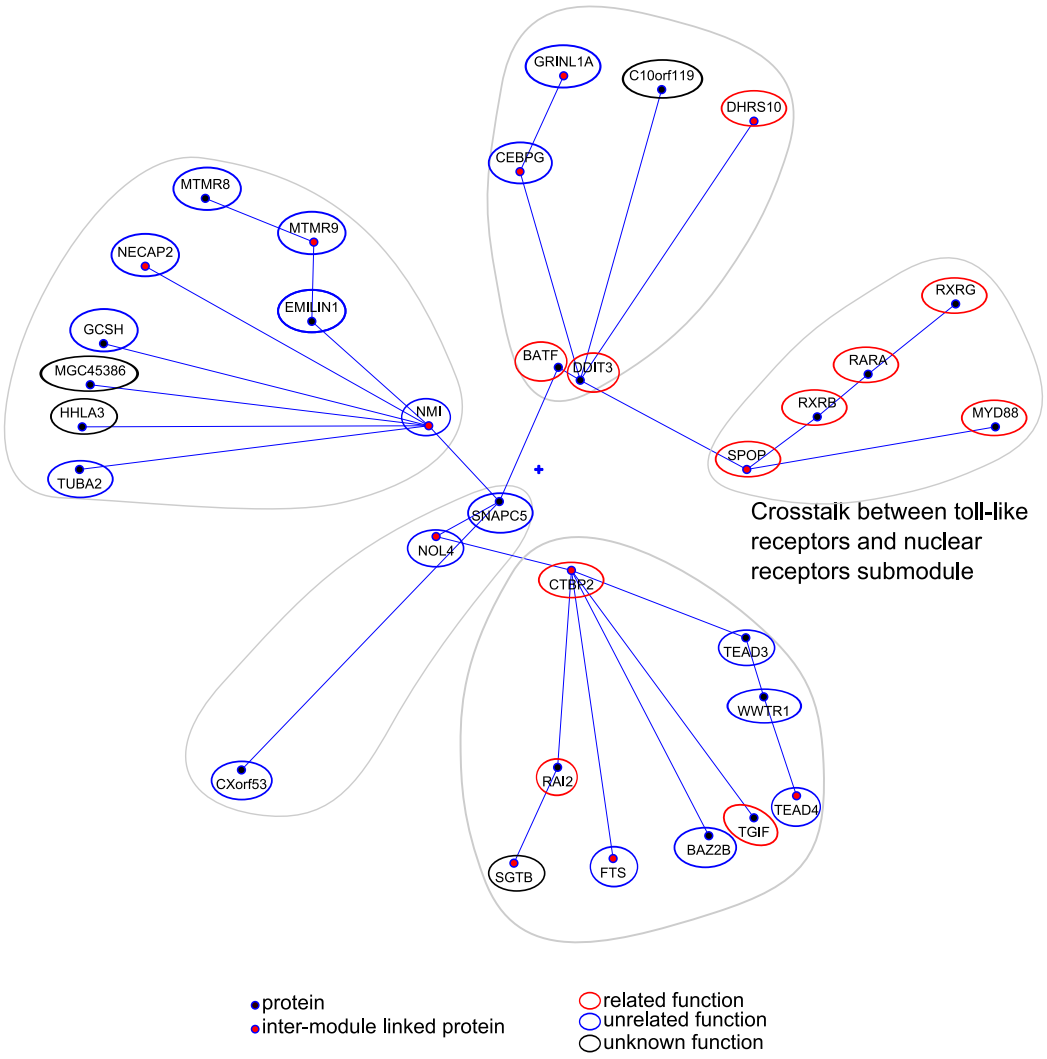


Figure 5: *Human Interactome module containing "Cross-talk between toll-like receptors (TLRs) and nuclear receptors" submodule.* Local Polar Map of module 2 in Figure 3. The highlighted submodule contains proteins involved in cross-talk between toll-like receptors and nuclear receptors. Note that two of the other submodules also contain some proteins in this category. See main text for a possible functional interpretation of this submodule. This figure was composed as described in the legend of Figure 4.

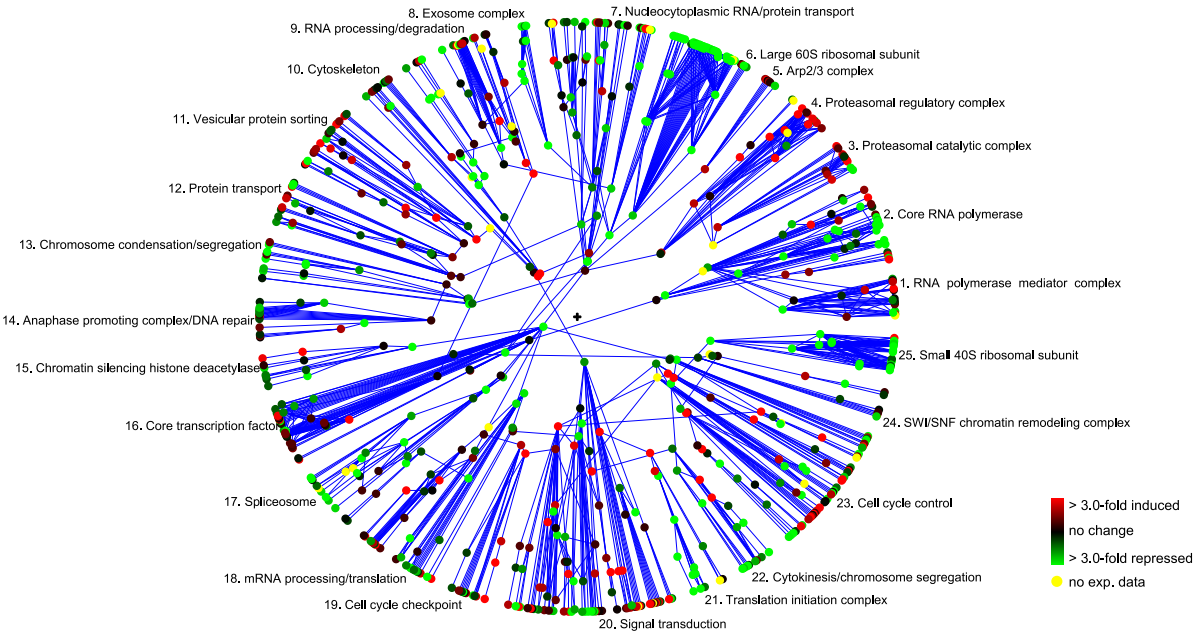


Figure 6: *Hydrogen peroxide oxidative stress gene expression response of* S. cerevisiae. Polar Map showing *S. cerevisiae* hydrogen peroxide oxidative stress mRNA expression data (0.30 mM during 20 min) from Gasch *et al.* [28] overlayed on the largest connected component of the yeast protein interaction network, based on the FYI dataset (741 proteins, 1752 interactions) [76] . Colors show mRNA fold induction relative to control unchallenged cells. See the main text for a description of this oxidative stress response.

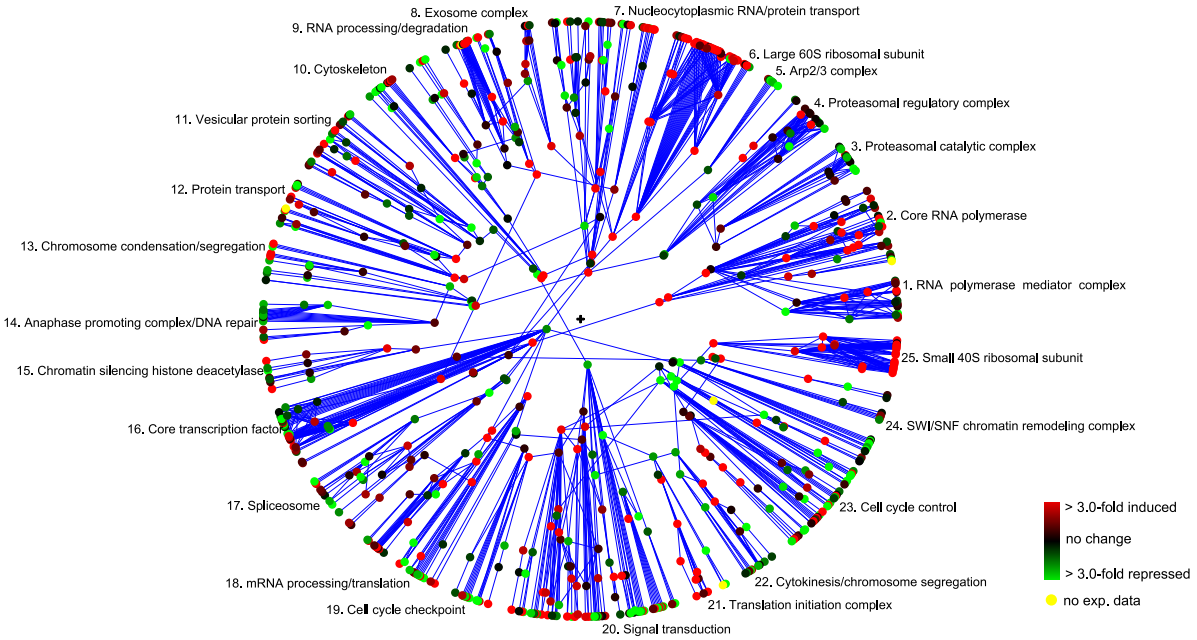


Figure 7: *Hydrogen peroxide oxidative stress vs. heat shock gene expression response in* S. cerevisiae. Analogous figure to Figure 6, except this time with colors showing hydrogen peroxide oxidative stress mRNA fold-induction *relative* to heat shock stress. See main text for analysis.