Table 1. Results on Human Y2H PPI network:

CATEGORIES	
final clusters	31
cover GRID complexes	3
$0 < PPIP < 17$	10
$18 < PPIP < 49$	3
$50 < PPIP < 89$	3
$90 < PPIP \leq 100$	3
NON-ZERO PPIP total	19
$> 50\%$ *PPIP total*	6
zero-PPIP +compatible location	9

Antonina Mitrofanova, Martin Farach-Colton and Bud Mishra

1 PROTEIN COMPLEXES OF OTHER SPECIES

We have applied our method to other species such as mouse, human, and worm. The Y2H interactions and protein complex information were obtained from the BIND database.

1.1 Human Protein Complexes

Each 2-edge connected human protein complex corresponds to one unique 2-edge connected subgraph in the Y2H network (2-edge connected subgraphs of existent 5 protein complexes showed no overlap). In our experiments, we identify 4 out of 5 2-edge connected protein complexes. Clusters that did not cover protein complexes were tested against compatible location evidence (cellular location information for every human protein was taken from swiss-prot (1) database). And indeed, all clusters responded positively to compatible cellular location tests and were not identified as false positives with respect to protein complexes.

However, we found an additional method for testing our predictions. We used hsPPIP (2) software to verify if clusters produced by the Homo Sapience Y2H PPI network are possible candidates for protein complexes. This software uses the idea that Protein Protein Interaction Predictions are similar in different organisms. It takes human proteins/genes of interest as input. These genes are compared to Yeast proteins/genes for possible orthologs. If such orthologs in Yeast participate in a protein complex, then the probability that human orthologs will form a complex is non-zero. This probability depends on the degree of orthology and the possibility of Yeast orthologous proteins to build a protein complex. Thus, when such predictions are positive, they can lend some support to the hypothesis that the cluster is a complex, whereas when the prediction is low, they yield no information.

The considered human Y2H PPI network consists of 2699 proteins and 3360 interactions. Figure 1 shows results of two runs ($d = 13$, 63 nodes eliminated) on the human Y2H PPI network with corresponding probabilities from hsPPIP (the higher is the probability, the better). All clusters can be viewed at http://research.rutgers.edu/$\sim$ amitrofa/predictions.html.

1.2 Worm Protein Complexes

The Worm Y2H PPI network consists of 3154 proteins connected by 4921 edges. It has one 2-edge connected protein complex, which was successfully identified by the algorithm ($d = 13$, 99 nodes eliminated). All other clusters were tested against compatible cellular location evidence, taken from (1). Unfortunately, the subcellular location information for Worm is very limited, to the extend that we were not able to obtain this information for more than one protein in a cluster. All 29 predicted protein complexes are can be found at http://research.rutgers.edu/$\sim$ amitrofa/predictions.html.

1.3 Mouse Protein Complexes

We also applied our method to protein networks that do not possess any known 2-edge connected protein complexes, such as the Mouse Y2H PPI network. We ran our algorithm with $d = 13$ (2 nodes eliminated) and identified 17 clusters, 13 of which responded to the compatible location evidence and 4 of which did not possess enough cellular location information (we did not find, however, any incompatible cellular locations among proteins in a single cluster). The Mouse network has 723 proteins and 630 edges. All mouse complex predictions can be found at http://research.rutgers.edu/$\sim$ amitrofa/predictions.html.

REFERENCES

[1]Swiss-prot: `http://ca.expasy.org/sprot/`
[2]shPPIP: $http://973 - proteinweb.ustc.edu.cn/hsppip/faq.php$