

Supplemental Table 1. Number of associations shared between pairs of protein interaction data sets. Assuming a complete interactome of 200,000 associations (i.e., ~10 unique associations each per 20,000 proteins), the 7,000 conserved co-expression associations show a statistically significant intersection ($p < 10^{-43}$) with the Stuart *et al.* dataset. Assuming a smaller overall interactome of 120,000 associations, no datasets significantly overlap the 7,000 set. Each comparison significant at 95% confidence ($p < 0.0018$ for 28 comparisons) is marked by an asterisk for the case of the complete interactome of size 200,000. Those significant assuming an interactome of 120K are marked by a dagger.

	Ramani et al	Stuart et al	Stelzl et al	Rual et al	7,000 conserved co-expression	Rhodes et al.	Ewing et al	OPHID
Ramani et al	31,609							
Stuart et al	1,620	15,000						
Stelzl et al	18	8	3,121					
Rual et al	95	5	9	2,754				
7,000 conserved co-expression	859	848*	2	1	7,000			
Rhodes et al.	2,718	441	14	59	195	38,379		
Ewing et al	128	13	1	9	15	56	4,670	
OPHID	8,118	686	1,829*†	1,646*†	211	2,286	2,461*†	57,507

Supplemental Table 2. Proportions of true physical associations measured for the CCE pairs and two control sets, using the methods of Figures 3 and 6 and an alternate gold standard positive association set derived from Ramani *et al.* (2005). Ranges of values are derived by comparison to the corresponding standard curves. Estimates of variance (+/- s.d.) for the path length method and ranges for the co-elution method are average values derived from analysis of control mixtures of known proportions of true and false positives.

	Percentage of true physical associations as measured by:					Average of five tests (+/- s.d.)
	Shotgun proteomics co-elution	Worm/fly/yeast physical interaction	GO/KEGG overlap	SwissProt keyword overlap	Yeast network path length	
7,000 conserved co-expression	49 – 59	53 – 64	51 – 58	59 – 68	76 +/- 3	61 +/- 9
Human-only co-expression (top 7,000)	18 – 28	1 – 2	1 – 3	12 – 22	25 +/- 3	14 +/- 11
7,000 randomized	0 – 5	4 – 10	2 – 4	3 – 8	5 +/- 3	5 +/- 2

Supplemental Table 3. Number of associations from the 7,000 conserved co-expression-derived association set whose strongest support (*i.e.*, maximum LLR score) derives from the given human-model organism datasets.

	<i>A. thaliana</i>	<i>C. elegans</i>	<i>D. melanogaster</i>	<i>M. musculus</i>	<i>S. cerevisiae</i>
Number of associations	1,079	2,949	1,244	158	1,570

Supplemental Table 4. mRNA expression data used for conserved co-expression analysis. All 3,977 DNA microarray experiment datasets were obtained from the Stanford Microarray Database. The 1,922 experiments corresponding to the human microarrays were split into 11 general categories for analysis (as described in Methods); data for other organisms was treated as single categories.

Organism	Category	Number of Experiments
Human	Non-tumor tissue	105
Human	Cell cycle	116
Human	Normal tissue	156
Human	Lymphoma	238
Human	Pathogen	330
Human	Primary tumor	527
Human	Breast cell line	58
Human	Tissue rejection	67
Human	T-cell	68
Human	Development	89
Human	Cell line	168
<i>A. thaliana</i>	-	574
<i>C. elegans</i>	-	553
<i>D. melanogaster</i>	-	170
<i>M. musculus</i>	-	113
<i>S. cerevisiae</i>	-	645

Supplemental Table 5. Number of orthologs (more specifically, in-paralogs (Remm et al., 2001)) included in the analysis of conserved co-expression, derived from the InParanoid database with subsequent mapping of SwissProt gene identifiers to the organism-specific identifiers.

	<i>A. thaliana</i>	<i>C. elegans</i>	<i>D. melanogaster</i>	<i>M. musculus</i>	<i>S. cerevisiae</i>
Human	5,820 4,393	5,867 6,773	5,898 7,570	10,357 10,378	2,142 3,513