Supplementary information

Tables:

Table S1. Details of the yeast PPI dataset used in our work.

DataSet	Number of Nodes	Number of Interactions	Platform
LC-Kim[1]	3268	12058	Literature-curated
IDBOS-Gavin[2]	1274	7879	AP/MS
IDBOS-Krogan[3]	1719	3640	AP/MS
Y2H-Union[4]	2019	2923	Y2H

1. Data is from [1]; 2. This data is from [2], we used the default cutoff 5.95; 3. This data is from [2], we used the default cutoff 12.92; 4. Data is from [3].

Table S2. The significance of the interactions within the age-homogeneous motifs.

dataset	motif size	number of age-homogenous motifs	deltaD after removing interactions within age-homogeneous motifs	original deltaD	p-value
LC-Kim	3	2212	0.11909275	0.54	$<10^{-3}$
	4	44562	-0.461184082		
IDBOS-Gavin	3	4644	-0.29278693	0.48	$<10^{-3}$
	4	39456	-0.314663703		
IDBOS-Krogan	3	1272	0.127818538	0.88	$<10^{-3}$
	4	3879	0.203066283		
Y2H-Union	3	40	0.5975674	0.7	$<10^{-3}$
	4	195	0.588700861		

Table S3. The number of motifs containing paralogs in age pattern #3 and 4#. The percentage is simply computed by the second column/third column to show a direct result.

Motif	Number of Motifs Containing Paralogs	Number of Motifs Containing Paralog PairS in Age Patterns 3# and 4#	Percentage	Dataset
B	69	24	34.78%	LC
E	509	46	9.03%	LC
G	72	8	11.11%	LC
H	16	3	18.75%	LC
B	50	7	14%	IDBOS-Gavin
E	289	14	4.84%	IDBOS-Gavin
G	52	4	7.69%	IDBOS-Gavin
H	6	0	0	IDBOS-Gavin
B	93	9	9.67%	IDBOS-Krogan
E	328	2	0.60%	IDBOS-Krogan

G	152	3	1.97%	IDBOS-Krogan
H	87	2	2.29%	IDBOS-Krogan

Table S4. Parameters for all the canonical models.

Model	Node Size	Parameters	Initial network
PA Model[1]	3000	connections of new node m_0=4	4 isolated nodes
Symmetric DD Model[2]	3000	Divergence prob p=0.6, mutation prob q=0.1	4 fully connected nodes
Asymmetric DD Model[3]	3000	Divergence prob p=0.587	4 fully connected nodes
CG Model[4]	3000	Newman algorithm to divide the communities	4 fully connected nodes

1. Preferential Attachment Model 2. Symmetric Duplication-Divergence Model 3. Asymmetric Duplication-Divergence Model 4. Crystal Growth Model
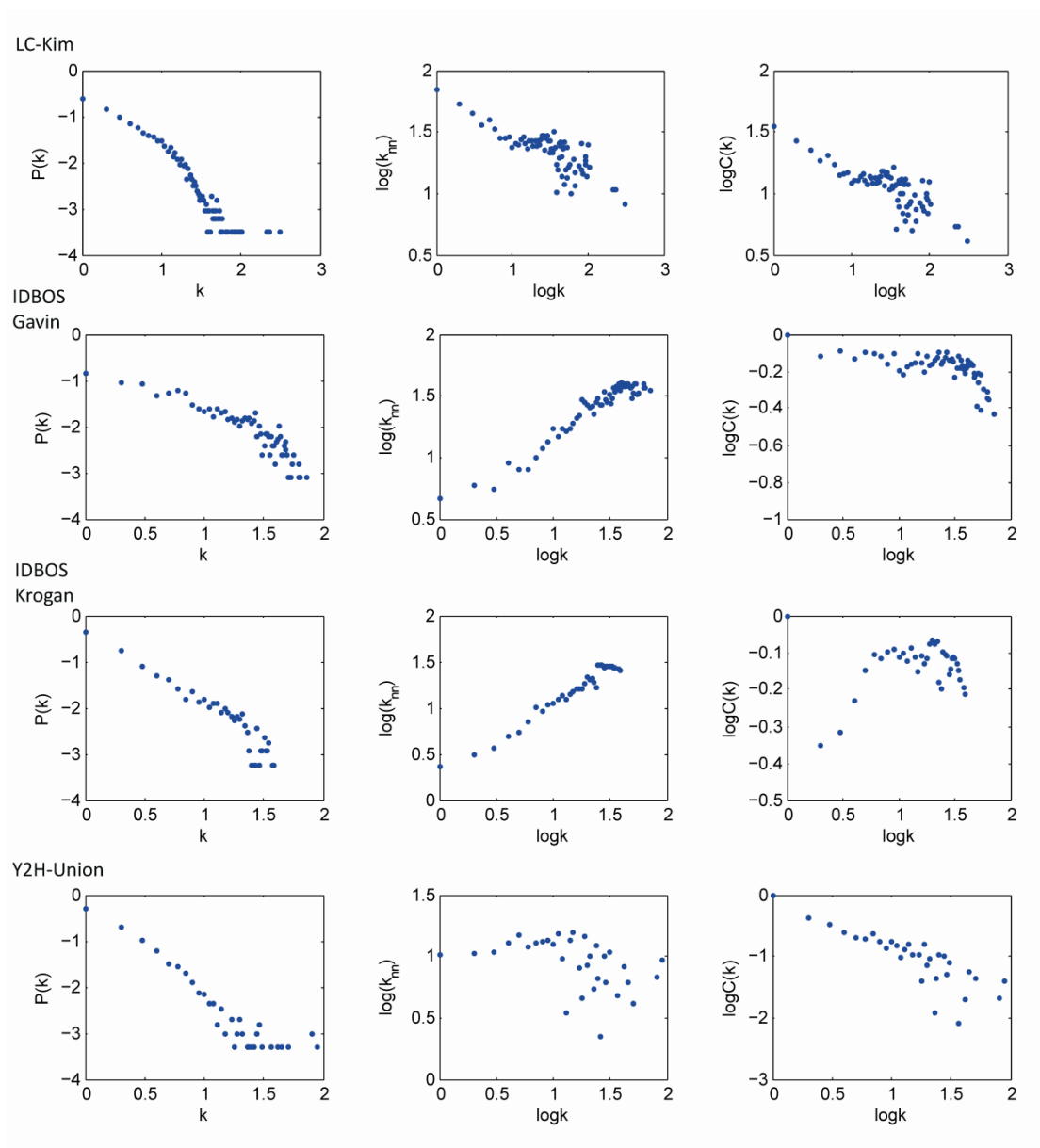
Figures:

Figure S1. The degree distribution P(k), clustering coefficient C(k) and the average degree of nearest neighbors k<nn> of the four yeast PPI networks.
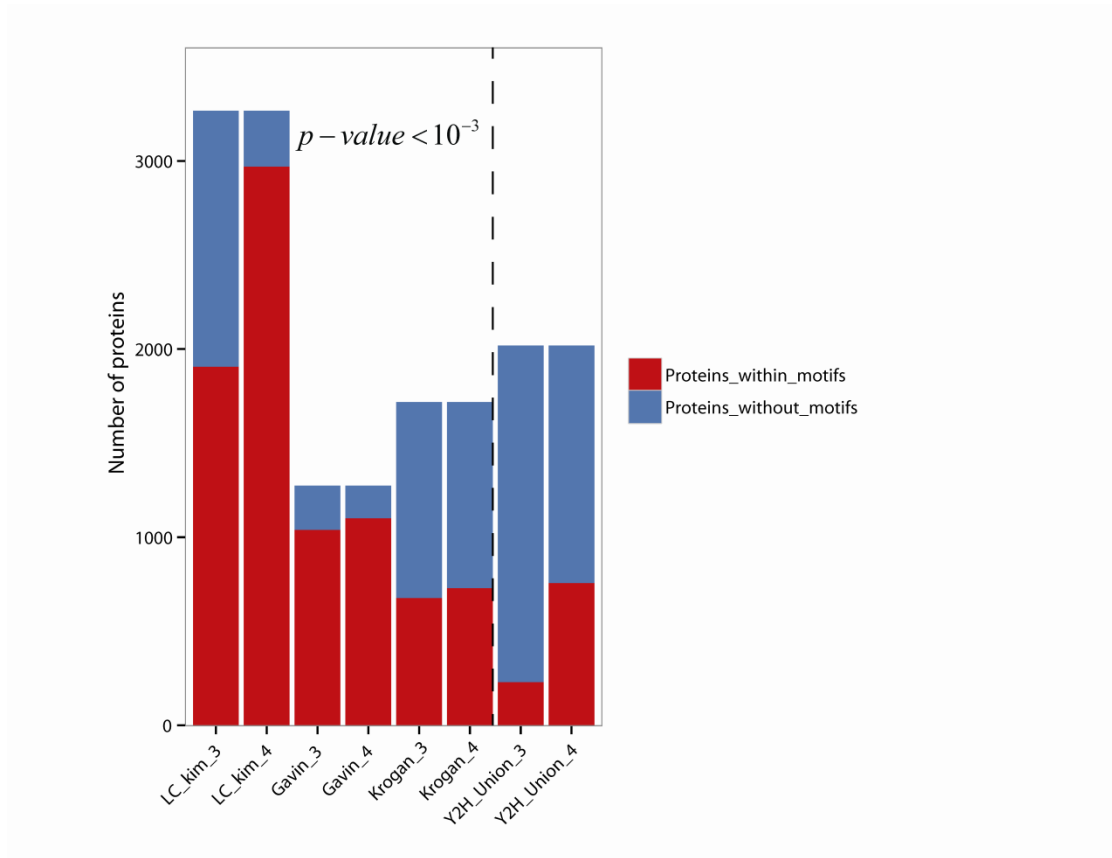
Figure S2. The number of proteins within/without network motifs across the four yeast PPI datasets. The p-value shows that proteins within network motifs among three datasets are significantly enriched in SGD_complexes where as no such tendency observed in Y2H-Union data. Same result holds for CYC_2008 copmlexes.
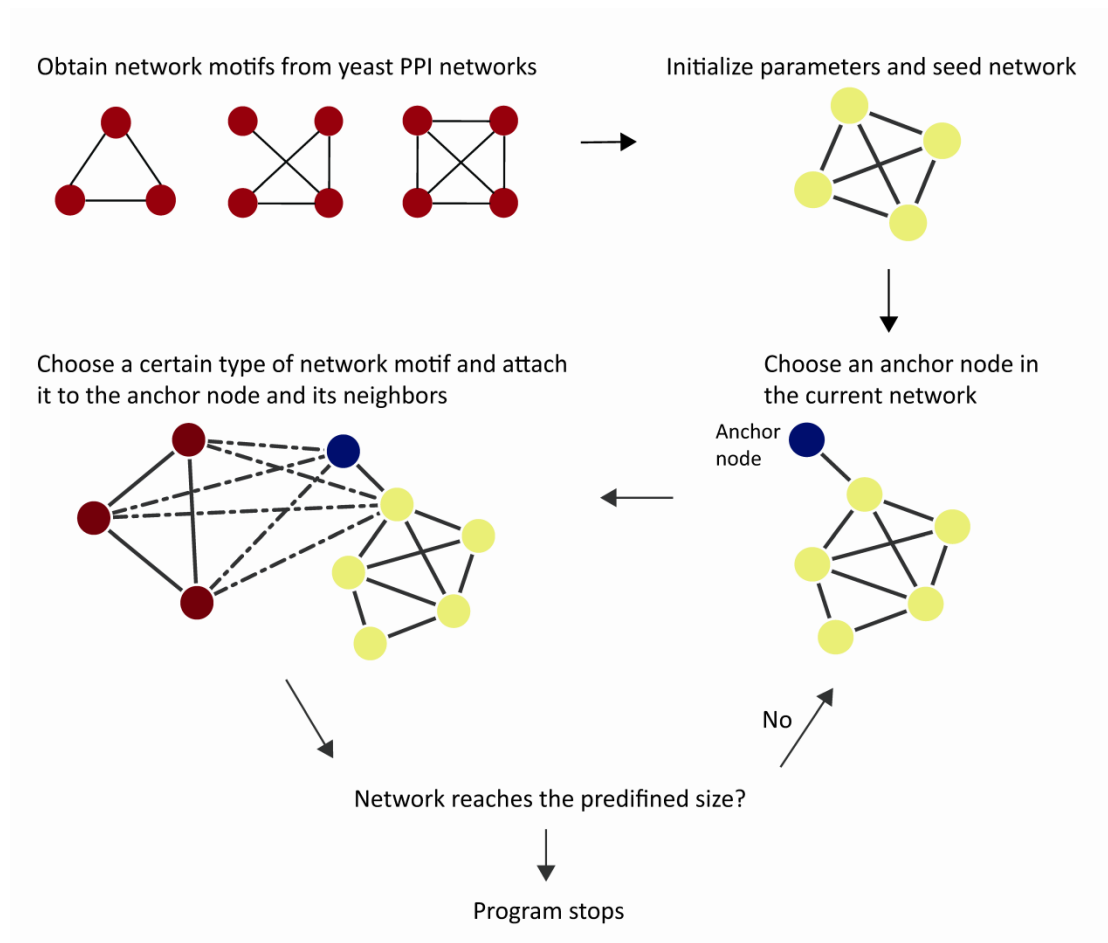
Figure S3. An illustration for our Network Motif model

1. Kim WK, Marcotte EM (2008) Age-dependent evolution of the yeast protein interaction network suggests a limited role of gene duplication and divergence. PLoS Comput Biol 4: e1000232.
2. Yu X, Ivanic J, Wallqvist A, Reifman J (2009) A novel scoring approach for protein co-purification data reveals high interaction specificity. PLoS Comput Biol 5: e1000515.
3. Yu H, Braun P, Yildirim MA, Lemmens I, Venkatesan K, et al. (2008) High-quality binary protein interaction map of the yeast interactome network. Science 322: 104-110.