Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 357 (2005) 501–512
PHYSICA A
www.elsevier.com/locate/physa

Graph theoretic analysis of protein interaction networks of eukaryotes

K.-I. Goh[a,*], B. Kahng[a,b], D. Kim[a]

[a]*School of Physics, Seoul National University, Seoul 151-747, Republic of Korea*
[b]*Program in Bioinformatics, Seoul National University, Seoul 151-747, Republic of Korea*

Received 2 March 2005
Available online 23 May 2005

Abstract

Owing to the recent progress in high-throughput experimental techniques, the datasets of large-scale protein interactions of prototypical multicellular species, the nematode worm *Caenorhabditis elegans* and the fruit fly *Drosophila melanogaster*, have been assayed. The datasets are obtained mainly by using the yeast hybrid method, which contains false-positive and false-negative simultaneously. Accordingly, while it is desirable to test such datasets through further wet experiments, here we invoke recent developed network theory to test such high-throughput datasets in a simple way. Based on the fact that the key biological processes indispensable to maintaining life are conserved across eukaryotic species, and the comparison of structural properties of the protein interaction networks (PINs) of the two species with those of the yeast PIN, we find that while the worm and yeast PIN datasets exhibit similar structural properties, the current fly dataset, though most comprehensively screened ever, does not reflect generic structural properties correctly as it is. The modularity is suppressed and the connectivity correlation is lacking. Addition of interologs to the current fly dataset increases the modularity and enhances the occurrence of triangular motifs as well. The connectivity correlation function of the fly, however, remains distinct under such interolog additions, for which we present a possible scenario through an in silico modeling.

Keywords: Protein interaction network; Modularity; Degree correlation

*Corresponding author.
E-mail address: vrooom@soeh.snu.ac.kr (K.-I. Goh).

0378-4371/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2005.03.044

1. Introduction

In the last few years, graph theoretic methods to understand complex biomolecular systems have been developed very rapidly [1]. Such a development has made advances toward uncovering the organizing principles of cellular networks in post-genomic biology. The cellular components such as genes, proteins, and other biological molecules, connected by all physiologically relevant interactions, form a full web-like molecular architecture in a cell. In such an architecture, genes which are expressed through proteins play a central role. Proteins rarely act alone, rather they cooperate with others to act physiologically. Thus, protein interactions play pivotal roles in various aspects of the structural and functional organizations and their complete description would be the first step toward a thorough understanding of the web of life. Proteins are viewed as nodes of a complex protein interaction network (PIN) in which two proteins are linked if they physically contact each other. The graph theoretic approach has been useful to understand intricate interwoven structures of the PIN [2–4]. The key biological processes indispensable to maintaining life are robust across eukaryotic species since many involved genes are evolutionarily conserved [5]. Using this property, one can test a newly discovered dataset if it really contains more or less complete information of protein interactions. Moreover, this in silico approach offers one the candidates of protein interaction pairs, of which the number is considerably reduced compared with the total combinatorial pairs. Thus, the graphic theoretic analysis would provide a useful guide for further wet studies of protein interactions.

Species with sequenced genome such as the yeast *Saccharomyces cerevisiae* provide important test beds for the study of the PIN. Owing to the recent progress in the high-throughput experimental techniques such as the yeast two-hybrid assay [6,7] and the mass spectroscopy [8,9], the dataset of the yeast PIN has been firmly established [10,11]. Very recently, large-scale protein interactions of multicellular species, the nematode worm *Caenorhabditis elegans* [12] and the fruit fly *Drosophila melanogaster* [13], have been assayed. While those datasets, mainly based on the yeast two-hybrid assay, need physiological proof, they contain large-scale proteins and protein interactions, making graph theoretic study possible. In this paper, we analyze those datasets and compare them with the more-established set of interactions in the budding yeast [11]. Our graph theoretic analysis suggests that the present interaction dataset of the fruit fly, based on the yeast two-hybrid (Y2H) assay, may have left out a significant part of protein interactions, though most comprehensively screened ever. Such a conclusion was reached by the comparison of the generic features of the PIN, the modularity and connectivity correlations, across the three species. For the fly, those quantities behave distinctively: The modularity is suppressed and the connectivity correlation is lacking. Such distinct behavior can be overcome partially by the addition of yeast interologs into the fly dataset.

2. Materials and methods

2.1. Graph theory terminology

Till recently, it has been well known that the number of connections of a given protein in the PIN is inhomogeneous and its distribution, called the degree

distribution, follows a power law. To characterize the structure of the PIN in a more concrete way, such a degree distribution is insufficient, and more graph theoretic quantities have been introduced. Here, we give the following quantities:

(i) Network is composed of vertices (nodes) and edges (links). In the PIN, vertices represent proteins and edges protein interactions.
(ii) Degree is the number of edges connected to a given vertex. The degree distribution $p_d(k)$ is the fraction of vertices having k degree.
(iii) Clustering coefficient of a vertex is defined as $C_i = 2e_i/k_i(k_i - 1)$, where e_i is the number of connections among the k_i neighbors of a vertex i. Clustering function $C(k)$ is the mean value of C_i over the vertices with degree k, while the clustering coefficient C is the mean of C_i over all vertices. When the network contains hierarchical and modular structures within it, it is known that the clustering function $C(k)$ behaves as $C(k) \sim k^{-\beta}$ for large k [14].
(iv) $\langle k_{nn} \rangle (k)$ is the mean degree of the neighbors of a vertex with degree k. It is known that $\langle k_{nn} \rangle (k) \sim k^{-\nu}$ with $\nu > 0$ for the Internet and the PIN [4,15], implying that vertices with large degrees tend to connect to the ones with small degree. Such a network is called disassortative network.
(v) The mixing coefficient r has been introduced [16] to characterize the degree–degree correlation between the two vertices located at the ends of an edge, which is defined as

$$r = \frac{\langle k_1 k_2 \rangle - \langle (k_1 + k_2)/2 \rangle^2}{\langle (k_1^2 + k_2^2)/2 \rangle - \langle (k_1 + k_2)/2 \rangle^2},$$

where k_1 and k_2 are the degrees of two vertices at the ends of an edge, and $\langle \cdots \rangle$ denotes the average over all edges.

2.2. The PIN datasets

We used the yeast subset of the interaction data compiled in the Database of Interacting Proteins (DIP) as of January 2004 (http://dip.doe-mbi.ucla.edu) [11]. The datasets for the worm and fly are obtained from the works of Li et al. [12] and Giot et al. [13], respectively. For the worm, we consider two different versions, the one consisting of only the interactions from the Y2H screens (referred to as Worm-Y2H network in this paper) and the other the full network supplied by Li et al. [12] (referred to as Worm-All network). The characteristics of each dataset and the values of the graphic theoretic quantities are shown in Table 1.

2.3. Orthologous gene assignment

For cross-species ortholog information, we used the information from the KOG database [17], a eukaryotic extension of the Clusters of Orthologous Genes (COG) database (http://www.ncbi.nlm.nih.gov/COG/new/).

Table 1
PIN datasets

	Yeast	Worm-Y2H	Worm-All	Fly
N_{proteome}	6195	22246	22246	16206
N	4714	2835	3216	7055
L	14857	4438	50444	20947
$\langle k \rangle$	6.3	3.1	3.4	5.9
C	0.12	0.047	0.15	0.014
r	−0.14	−0.16	−0.13	−0.036
N_1	4627	2601	2898	6929

The size of proteome N_{proteome}, the number of proteins N and the number of protein–protein interactions L in the dataset, the mean degree $\langle k \rangle$, the clustering coefficient C, the mixing coefficient r, and the number of proteins forming the largest cluster N_1 are tabulated for each dataset. The self-interactions are excluded throughout.

2.4. Yeast interologs in fly

Having identified the yeast–fly orthologs [18], we look for the interactions in the yeast network between those yeast proteins both having orthologs in the fly network. Such orthologous interactions are called the interologs. If the corresponding fly interaction is present, we call it an *overlap interolog*. If not, we call it a *potential interolog*. Note that the ortholog relationship is not always one-to-one, resulting in multiple interologs for a given yeast interaction. For in silico analysis on the effect of the addition of potential interologs in the fly network, we include on average one potential interolog per yeast interaction. Specifically, for each yeast interaction A-B having no overlap interolog, each potential interolog is added in the fly network with equal probability, that is, $1/(o_A o_B)$, where o_X is the number of fly ortholog(s) of the yeast gene X. The network obtained in this way is referred to as Fly + Interolog network hereafter. The full list of the 408 overlap and the 55176 potential interologs are available on the web (http://komplex0.snu.ac.kr/pin/yeast-fly-interolog.xls).

3. Results

3.1. Degree distributions

In Fig. 1, we plot the degree distributions of diverse PINs, all of which display the scale-free behavior, fitting well to the generalized Pareto formula, $p_d(k) \sim (k + k_0)^{-\gamma}$, almost indistinguishable with each other. While the degree distribution is a fundamental quantity in graph theory, it deals with global network structure, so it does not give detailed information on structural property.

3.2. Modularity

A cellular function is achieved by a set of related proteins, usually forming a pathway or a complex. Such a functional module manifests itself as a localized dense

subgraph within the whole cellular network. The presence of modules and their hierarchical organization can be visualized by the local clustering function $C(k)$ [14]. For the yeast PIN, $C(k)$ exhibits a plateau for small k and falls off rapidly for large k, reflecting the modular structure bridged by the hubs (Fig. 2a). A similar pattern is observed in the worm (Fig. 2c). Note that the worm dataset contains the yeast interologs. Such a behavior can be also observed in the two prokaryotic species, *H. pylori* and *E. coli* (Fig. 2b). For the fly Y2H data, however, $C(k)$ behaves distinctively, almost constant for all k (Fig. 2e). To understand this discrepancy, we add the potential yeast interologs into the current fly Y2H dataset. Then $C(k)$ behaves in a similar fashion to other datasets, showing a moderate plateau for small k and rapid decrease for large k, albeit the altitude of the plateau, which is roughly the clustering coefficient C, is not as high as in the yeast and the worm (Fig. 2f). To find the role of the interologs in the worm, we consider the Worm-Y2H dataset, and plot its $C(k)$ in Fig. 2d. Indeed, compared with Fig. 2c, the character of $C(k)$ is changed, in particular, the plateau for small k almost disappears, implying the yeast-interologs play a role of forming modules, where proteins are closely linked to each other.

3.3. Conservation rate of interactions

We count how many yeast interactions are actually conserved in orthologous form in both the worm and the fly. The conservation rate found in this way for the Y2H

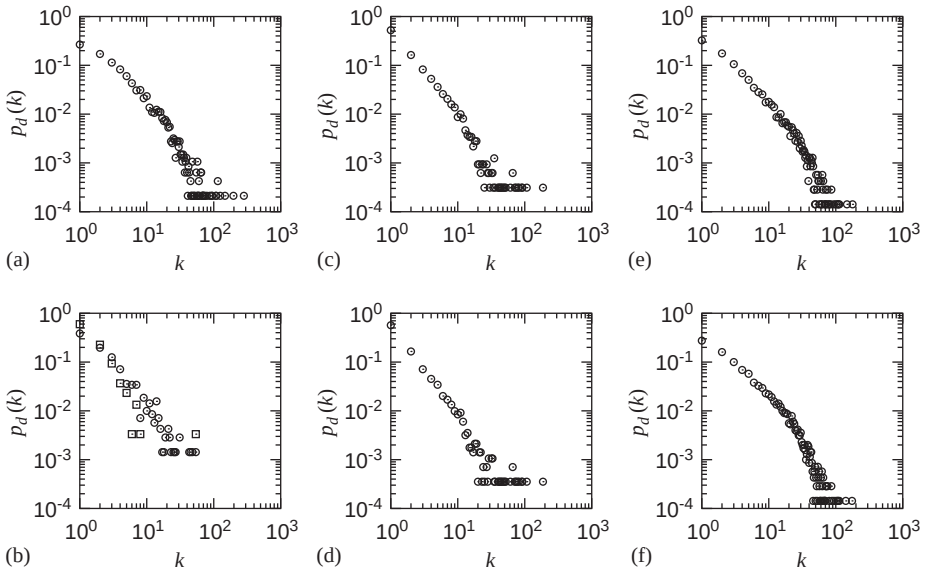

Fig. 1. The degree distributions $p_d(k)$ for (a) the yeast, (b) the prokaryotes *Helicobacter pylori* (◯) and *Escherichia coli* (□), (c) the worm (Worm-All), (d) the Y2H subset of the worm dataset (Worm-Y2H), (e) the fly, and (f) the Fly + Interolog dataset.

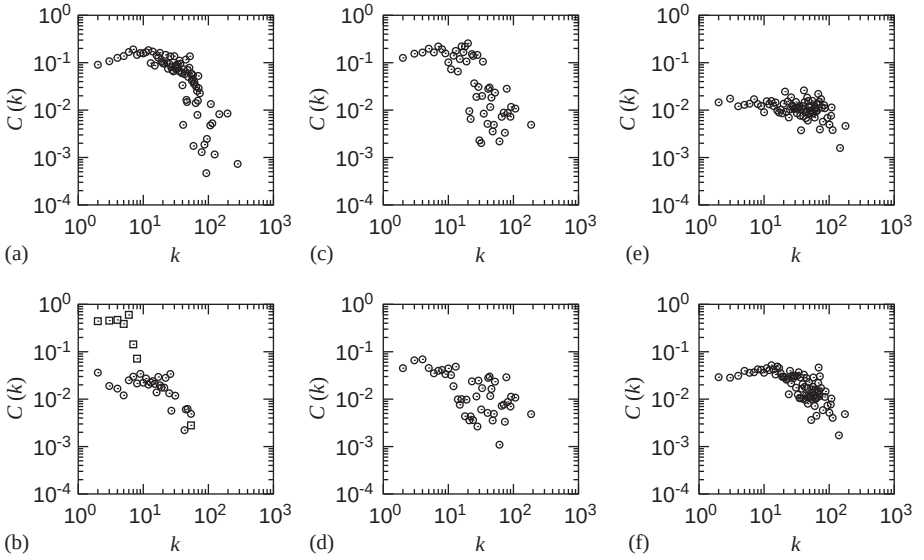

Fig. 2. The local clustering function $C(k)$ for (a) the yeast, (b) the bacteria *H. pylori* (◯) and *E. coli* (□), (c) the worm (Worm-All), (d) the Worm-Y2H dataset, (e) the fly, and (f) the Fly + Interolog dataset. The abscissae and ordinates are fixed for clear comparison.

screen dataset is surprisingly low; 2.7% for the worm (Worm-Y2H) and 3.8% for the fly. For the worm, we note that such low coverage is in part due to the insufficient number of baits used in the experiment (3,024 baits, 833 of which are present in the network). When we consider the conservation of triangular interaction patterns, a basic unit of cooperative functional module [19], only 3 out of 1731 are conserved in the worm, while none in the fly (Fig. 3). The lack of conserved interaction motifs in the fly data suggests that the current fly network misses some of the important cooperative aspects of the cellular network in the fly. The effort to fill this gap is required in time.

3.4. Motif structure

Since the modularity manifested by $C(k)$ is closely related to the formation of triangles in the network, here we further perform network motif analysis for the three species datasets. The network motifs are small recurring subgraphs which are overrepresented in a given network and are believed to provide the basic evolutionary and functional signatures of the network [19]. Since it was recently discovered that the motif constituents are more conserved during evolution than the rest [20], one would expect the density of each motif to be close to each other across the three species. From the comparison of the columns for Yeast, Worm-All, and Fly in Table 2, we can see that the triangular motif is relatively not abundant in Fly, while the square motif is. Thus, the absolute magnitude of the clustering function is

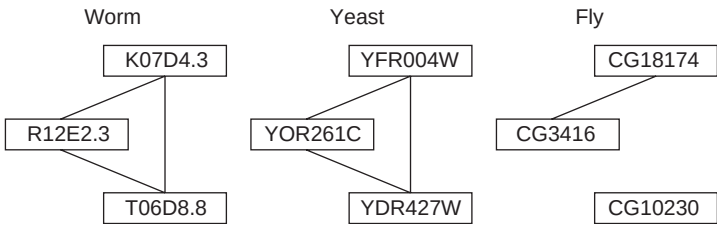

Fig. 3. Conservation of interaction motif. Shown in the middle is a triangular interaction subgraph within the yeast involving in ubiquitin-dependent protein catabolism. Corresponding orthologous counterparts in the worm and fly are also shown. This motif is conserved in the worm Y2H data, while only a single interaction is detected in the fly data.

smaller for the fly than for the yeast or the worm. The density of the triangular motif is higher in the Fly + Interolog dataset, indicating that the clustering coefficient is enhanced overall by the addition of the interologs of the fly.

In Table 2, we have summarized the motif structure for each network. We follow Milo et al. [19] to calculate the two scores, Z- and E-score, defined as $Z = (N - N_{random})/\sigma_{random}$ and $E = (N - N_{random})/N_{random}$, respectively, and use the following two criteria to specify whether a subgraph is a motif or an anti-motif (an anti-motif is a subgraph significantly underrepresented in the network):

(i) The probability that N is observed in randomized network is smaller than 0.01.
(ii) $|E| > E_0$, where we set the threshold $E_0 = 0.5$, rather than $E_0 = 0.1$ in Milo et al. [19].

Here, N_{random} and σ_{random} are the expected numbers of occurrence in the randomized version of the network and their standard deviation obtained from 1000 samples, respectively, where the randomization is performed by the switching method [19]. In calculating them for the four-node subgraphs, the numbers of three-node subgraphs are fixed to be those of the original networks. For the Fly + Interolog network, ten realizations of interolog addition (see Method) are averaged.

3.5. Degree–degree correlation

The mean neighbor degree function $\langle k_{nn}\rangle(k)$is useful in understanding the degree–degree correlation in a network. In Fig. 4, we plot $\langle k_{nn}\rangle(k)$for each dataset. For the yeast, it is known that $\langle k_{nn}\rangle(k)$decreases with increasing k [4], which turns out to be also true for some prokaryotic species (Figs. 4a,b). Such a behavior in $\langle k_{nn}\rangle(k)$is also observed for the worm (Figs. 4c,d); however, it is flat for the fly, implying lack of correlation (Fig. 4e). Such a distinct behavior for the fly is robust under the addition of the interologs (Fig. 4f), which suggests that the lack of correlation in the fly network could be intrinsic, even though we cannot exclude the possibility that it is again the artifact of the incompleteness of the data. The hypothesis that the lack of correlation could be intrinsic may be supported by the following observations.

Table 2
Network motif structure of the three species

	Yeast	Worm-Y2H	Worm-All	Fly	Fly + Interolog
	329961	81205	87294	413926	520704 ± 1358
	7136 M ($Z = 80,\ E = 3.3$)	366	1512 M ($Z = 29,\ E = 2.5$)	1549	3504 ± 40 M ($Z = 45, E = 1.4$)
	4081023	604723	680485	7378808	971960 ± 37157
	9024723	2129609	2157048	6315922	7409320 ± 24476
	368730 AM ($Z = -122,\ E = -0.7$)	46050	58520 AM ($Z = -59,\ E = -0.7$)	160846	263324 ± 2617
	21806	4350 M ($Z = 9.5,\ E = 0.6$)	4686	54100 M ($Z = 81,\ E = 2.1$)	60648 ± 206 M ($Z = 60,\ E = 1.3$)
	27455 AM ($Z = -49,\ E = -0.7$)	1505	4120 AM ($Z = -25,\ E = -0.6$)	4029 M ($Z = 12,\ N = 0.8$)	9313 ± 228
	5259	30	1563 M ($Z = 10,\ E = 0.8$)	82 M ($Z = 11,\ E = 3.5$)	914 ± 35 M ($Z = 40,\ E = 6.0$)

The number of each subgraph present in the network is tabulated. According to its Z- and E-score, the significant motifs (M) and anti-motifs (AM) are indicated.

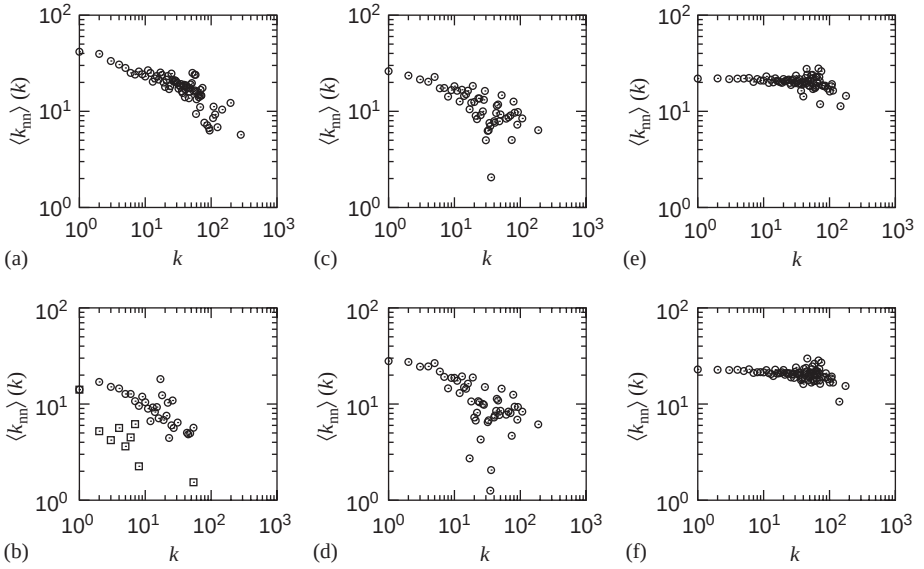

Fig. 4. The average neighbor degree function $\langle k_{nn}\rangle(k)$ for (a) the yeast, (b) the prokaryotes *H. pylori* (◯) and *E. coli* (□), (c) the worm (Worm-All), (d) the Y2H subset of the worm (Worm-Y2H), (e) the fly, and (f) the Fly + Interolog dataset. The abscissae and ordinates are fixed for clear comparison.

3.6. *Effect of diversification of gene function on* $\langle k_{nn}\rangle(k)$

While the pattern of $C(k)$ of the fly becomes similar to those of the yeast and the worm by the addition of the interologs, that of $\langle k_{nn}\rangle(k)$ remains distinct. Thus, here we investigate if such a flat behavior is intrinsic through an in silico model, finding that indeed, the decreasing behavior of $\langle k_{nn}\rangle(k)$ becomes moderated through the network evolution with the duplication and divergence processes. Homologs in a genome are thought to result from the gene duplication event, which is usually followed by the diversification to lower the redundancy. Some computer models aiming to mimic these processes in proteome evolution exist in the literature [21,22]. We investigate how the diversification process affects the topological property of the proteome network, in particular, the degree–degree correlation in terms of $\langle k_{nn}\rangle(k)$. To this end, we perform the following procedures motivated by Vázquez et al. [22]:

(1) Starting with the yeast protein network, at each step, a protein A is chosen randomly and is duplicated as A′. Then the protein A and A′ share common neighbors.
(2) For each neighboring protein of A and A′, one of edges connected to either A or A′ is removed with equal probability.
(3) Repeat 1 and 2 until the number of proteins reaches $\sim 20,000$, the approximate sizes of the worm and the fly proteome.

Note that in this procedure, the number of proteins increases while the number of interactions stays still. Thus, the average degree decreases as the size of proteome increases. Such a decrease will be compensated by, e.g., the acquisition of new interactions between existing proteins via mutation. However, we do not take such a process into account, to single out the effect of the diversification only.

The result of simulation is shown in Fig. 5. The local clustering function $C(k)$ is simply shifted downward, due to the overall decrease of the edge density. On the other hand, the average neighbor degree $\langle k_{\rm nn}\rangle(k)$decreases as k but with a smaller rate, indicating that the diversification process can, although not perfectly, neutralize the connectivity correlation. Furthermore, if we *assume* that the establishment of new interactions follows the preferential attachment [23] or random attachment, the overall correlation would diminish eventually.

3.7. *Effect of bait selection on* $\langle k_{\rm nn}\rangle(k)$

There has been an argument that the apparent decreasing trend in $\langle k_{\rm nn}\rangle(k)$ is an artifact from the limited selection of baits in the two-hybrid experiment [24]. Indeed, Li et al. [12] had selected the baits with their own criteria, mainly based on the biological indispensability and the potential applicability to the human therapeutics. To check this hypothesis in silico, we sampled the 30% subset of 4950 baits identified in Giot et al.'s fly network [13] and reconstructed the network only with the interactions associated with the sampled baits. We sampled in two different ways; the random sampling and the biased sampling toward the highly connected baits (the sampling probability is proportional to the number of bait-interactions). Both data sets generate the decreasing trend in $\langle k_{\rm nn}\rangle(k)$ (Fig. 6). One can see that even though the original network has the null slope in $\langle k_{\rm nn}\rangle(k)$, the negative slope develops in the sampled ones, demonstrating that the insufficient use of the bait *can* produce artifactual correlation in the connectivity. If this scenario holds, one

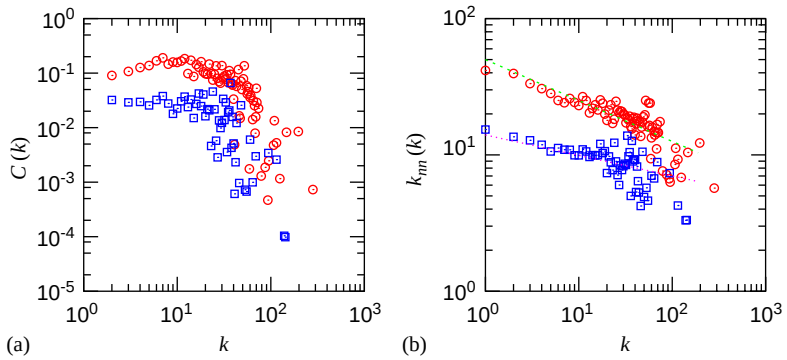

Fig. 5. Effect of gene function diversification in (a) $C(k)$ and (b) $\langle k_{\rm nn}\rangle(k)$. Red circles are the data of the original yeast network and the blue squares those after running the diversification procedures in silico. The slope of the straight line (the rate of decrease) in (b) is -0.3 (top, green) and -0.15 (bottom, magenta), respectively.

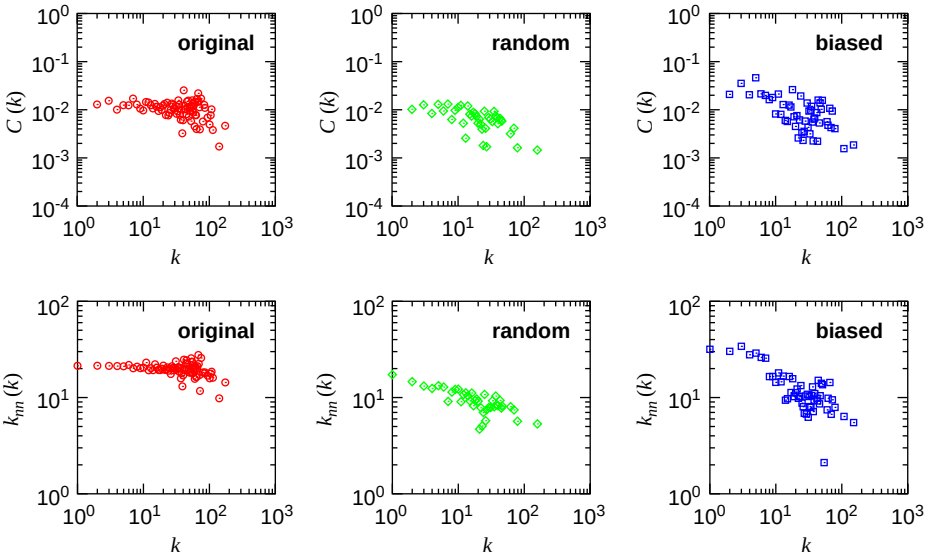

Fig. 6. Effect of bait selection. Red circle is for the full data, green diamond for the randomly sampled one, and the blue square for the biased sampled one.

conjecture is that the $\langle k_{\rm nn}\rangle(k)$ curve will become flatter as the interaction data accumulates and becomes more complete.

4. Summary and discussion

We have investigated in detail the structural properties of the protein interaction networks of three eukaryotic species, the budding yeast, the nematode worm, and the fruit fly. In particular, we have focused on the comparative assessment of the modularity and the degree–degree correlation for those networks. We found that while the worm dataset behaves similar to the yeast for the two graph theoretic quantities, the fly does not. The difference might be attributed to the presence (absence) of the yeast-interologs in the current worm (fly) dataset. For the fly dataset, the modularity is suppressed and the connectivity correlation is lacking. We found that the clustering function can be restored to those of the yeast dataset by the addition of interologs selected randomly among the candidates to the current dataset. We also performed motif analysis for the three species, finding that the density of the triangle motif is increased by the addition of the interologs to the current fly dataset. Finally, the candidates of the protein interactions of the fly are provided in the supplementary materials, which could be useful in finding protein interactions missed in the current fly dataset.

Acknowledgments

This work was supported by the KOSEF Grant no. R14-2002-059-01000-0 in the ABRL program and the MOST Grant no. M1 03B500000110.

References

[1] A.-L. Barabási, Z.N. Oltvai, Nat. Rev. Genet. 5 (2004) 101.
[2] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, Nature 411 (2001) 41.
[3] A. Wagner, Mol. Biol. Evol. 18 (2001) 1283.
[4] S. Maslov, K. Sneppen, Science 296 (2002) 910.
[5] B. Alberts, D. Bray, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, Essential Cell Biology, Garland, New York, 1998.
[6] P. Uetz, L. Giot, G. Cagney, T.A. Masfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, et al., Nature 403 (2000) 623.
[7] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, Y. Sakaki, Proc. Natl. Acad. Sci. USA 98 (2001) 4569.
[8] A.-C. Gavin, M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat, et al., Nature 415 (2002) 141.
[9] Y. Ho, A. Gruhler, A. Heilbut, G.D. Bader, L. Moore, S.L. Adams, A. Millar, P. Taylor, K. Bennett, K. Boutilier, et al., Nature 415 (2002) 180.
[10] H.W. Mewes, C. Amid, R. Arnold, D. Frishman, U. Guldener, G. Mannhaupt, M. Munsterkotter, P. Pagel, N. Strack, V. Stumpflen, et al., Nucleic Acids Res. 32 (2004) D41.
[11] L. Salwinski, C.S. Miller, A.J. Smith, F.K. Pettit, J.U. Bowie, D. Eisenberg, Nucleic Acids Res. 32 (2004) D449.
[12] S. Li, C.M. Armstrong, N. Bertin, H. Ge, S. Milstein, M. Boxem, P.O. Vidalain, J.D. Han, A. Chesneau, T. Hao, et al., Science 303 (2004) 540.
[13] L. Giot, J.S. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y.L. Hao, C.E. Ooi, B. Godwin, E. Vitols, et al., Science 302 (2003) 1727.
[14] E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltvai, A.-L. Barabási, Science 297 (2002) 1551.
[15] R. Pastor-Satorras, A. Vázquez, A. Vespignani, Phys. Rev. Lett. 87 (2002) 258701.
[16] M.E.J. Newman, Phys. Rev. Lett. 89 (2002) 208701.
[17] R.L. Tatusov, N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S.L. Mekhedov, A.N. Mikolskaya, et al., BMC Bioinformatics 4 (2003) 41.
[18] Ortholog means the gene or protein in different species but derived from a common ancestral gene.
[19] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Science 298 (2002) 824.
[20] S. Wuchty, Z.N. Oltvai, A.-L. Barabási, Nat. Genet. 35 (2003) 176.
[21] R.V. Solé, R. Pastor-Satorras, E. Smith, T. Kepler, Adv. Complex Syst. 5 (2002) 43.
[22] A. Vázquez, A. Flammini, A. Maritan, A. Vespignani, Complexus 1 (2003) 38.
[23] A.-L. Barabási, R. Albert, Science 286 (1999) 509.
[24] P. Aloy, P.B. Russell, FEBS Lett. 530 (2002) 253.