

## Rich can get poor: conversion of hub to non-hub proteins

Kyaw Tun · Raghuraj Keshava Rao ·  
Lakshminarayanan Samavedham · Hiroshi Tanaka ·  
Pawan K. Dhar

Received: 20 February 2009 / Revised: 6 April 2009 / Accepted: 7 April 2009 / Published online: 28 April 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** Hubs are ubiquitous network elements with high connectivity. One of the common observations about hub proteins is their preferential attachment leading to scale-free network topology. Here we examine the question: does rich protein always get richer, or can it get poor too? To answer this question, we compared similar and well-annotated hub proteins in six organisms, from prokaryotes to eukaryotes. Our findings indicate that hub proteins retain, gain or lose connectivity based on the context. Furthermore, the loss or gain of connectivity appears to correlate with the functional role of the protein in a given system.

**Keywords** Hub · Non-hub · Proteins · Rich · Poor · Connectivity

### Abbreviations

$k$	Connectivity
$\langle k \rangle$	Average connectivity
$k_{fc}$	Connectivity fold change
$p(k)$	Probability distribution of connectivity
SS	Sum square error
df	Degree of freedom
PPI	Protein–protein interaction network

---

K. Tun · P. K. Dhar (✉)  
Synthetic Biology Lab, RIKEN Advanced Sciences Institute,  
1-7-22, Suehiro-cho, Tsurumi 230-0045, Yokohama, Japan  
e-mail: pkdhar@riken.jp

K. Tun  
e-mail: kyawtun@gsc.riken.jp

R. K. Rao · L. Samavedham  
Department of Chemical and Biomolecular Engineering,  
Engineering Drive, National University of Singapore,  
Singapore 117576, Singapore

R. K. Rao  
e-mail: raghuraj@nus.edu.sg

L. Samavedham  
e-mail: chels@nus.edu.sg

K. Tun · H. Tanaka  
Department of Systems Biology, School of Biomedical Sciences,  
Tokyo Medical and Dental University, Yushima, 1-5-45,  
Bunkyo 113-8510, Tokyo, Japan

H. Tanaka  
e-mail: tanaka@cim.tmd.ac.jp

### Introduction

In a network, hubs represent nodes with large number of links surrounded by nodes with just few connections. The hubs have important roles, not only in information management within a network but also as regulatory molecules (Rodriguez-Caso et al. 2005). Protein–Protein Interaction (PPI) networks, in their general form, appear as a small number of hubs and a large number of sparsely connected non-hub proteins. Recent evidence points to the preponderance of structural disorderliness in hub proteins compared with the non-hub proteins (Haynes et al. 2006; Singh et al. 2007).

A notable feature of the PPI networks is their property of power-law and scale-freeness. The differential degree of connectivity distribution has been ascribed to the preferential attachment of hub proteins (Barabasi and Albert 1999; (Barabasi and Oltvai 2004). According to this model, the probability of connecting to a new node in the network is proportional to the connectivity density (i.e., vertex degree) of the existing node. Thus, a highly connected node

has greater chances of attracting a new node compared to the ‘connectivity potential’ of a sparsely connected node, i.e., rich gets richer. Interestingly, enzymes that are candidates for horizontal gene transfer seem to have higher average connectivity than other enzymes (Light et al. 2005). Furthermore, protein age seems to correlate with the network connectivity (Eisenberg and Levanon 2003). Saturation and viability have also been reported as properties of networks showing preferential attachment (D’Souza et al. 2007). Though preferential attachment of hub proteins has been studied in the past, the reasons of its origin and the functional implications of ‘rich-gets-richer’ are unclear. Interestingly, Pagel et al. (2007) propose a variable rate of attachment model according to which power-scaling can be achieved without preferential attachment.

In this study we asked: Is preferential attachment universal to all the protein-protein interaction networks? Are there any exceptions? What is the functional implication of the preferential attachment model, or the lack of it? To answer these questions, well-annotated and evolutionary conserved hub orthologs in several organisms were identified. Six protein-protein interaction (PPI) databases were integrated to minimize the effect of data bias. Stringent criteria were followed to ensure the accuracy of the sample. Overall, our work supports the existing notion of “rich-gets-richer”. However, every hub protein that we studied did not follow the preferential attachment model. Several examples of ‘rich-proteins-getting-poor’ were observed. Based on these findings, we propose a hypothesis of *hub convertibility*, i.e., a hub protein can get rich or poor, based on the context. Our findings also point to the functional correlation with the loss or gain of hubness.

## Materials and methods

Six organisms, *H. pylori*, *E. coli*, *S. cerevisiae* (Yeast), *C. elegans* (Worm), *D. melanogaster* (Fly) and *H. sapiens* (Human) were selected in the present study due to their rich annotation and interaction data. The publicly available protein-protein interaction databases (DIP, Salwinski et al. 2004; BIND, Bader et al. 2003; IntAct, Hermjakob et al. 2004; Reactome, Joshi-Tope et al. 2005; HPRD, Mishra

et al. 2006; MINT, Chatr-aryamontri et al. 2007) were integrated. We used exact sequence homology to overcome the problem of unique accession identifier in each database. Using this meta-dataset, a sample of non-redundant proteins and their interactions was extracted (Table 1). In all, 44,149 proteins were collected from 100,915 accessions and 205,835 interactions were collected from 260,449 accessions. The average overlap for proteins was 1.95 and the average overlap for interaction between them is 1.26.

The InParanoid algorithm (O’Brien et al. 2005) with BLOSUM45 substitution matrix for prokaryote and BLOSUM60 for eukaryote was used to identify orthologs. The pair wise orthologs were combined into multi-species cluster using MultiParanoid algorithm (Alexeyenko et al. 2006). To separate orthologs from spurious matches, a 50 bits cut-off was used. The matching segment of the longer sequence exceeded 50% of its total length. The ortholog data were also examined using COG and KOG protein family classification.

We selected fold change definition for connectivity,  $k_{fc} = k/\langle k \rangle$ , and appropriate cutoff to identify hubs in different PPI networks. For Prokaryotes, a node with  $k_{fc} \geq 2$  was considered as hub (cutoff,  $P < 0.03$  using distribution of standard normalized kfc values in Ortho\_Pk). For eukaryotes the criterion was,  $k_{fc} \geq 10$  (with  $P < 0.001$ ). A higher cutoff of 10-fold change was used for eukaryotes in order to minimize the effect of false positives, especially in data from *S. cerevisiae*. Of the orthologs identified in each category (Ortho\_PkEk, Ortho\_Pk and Ortho\_Ek), only those satisfying stringent hub criteria in at least one species were selected. A hub was considered ‘converted to non-hub’ when its kfc value crossed the cutoff. In order to define the nature of this hub convertibility, we observed trends along the complexity profile of six species (*H. pylori*, *E. coli*, *S. cerevisiae*, *C. elegans*, *D. melanogaster* and *Homo sapiens*).

Three classes of hubs were identified from the data: (a) the “getting rich” hubs showing increasing  $k_{fc}$ , (b) the “getting poor” hubs showing decreasing  $k_{fc}$  and (c) “flexible” hubs with non-uniform connectivity trend across the organisms. The GO annotation for each protein was obtained from the source database (BIND, SGD, Flybase, Wormbase, HPRD). The resulting GO annotation are

**Table 1** PPI networks selected in this study

Item	<i>H. pylori</i> (HPY)	<i>E. coli</i> (ECL)	<i>S. cerevisiae</i> (YST)	<i>C. elegans</i> (WRM)	<i>D. melanogaster</i> (FLY)	<i>H. sapiens</i> (HMN)
Proteins (# of nodes)	758	3393	6498	7390	10363	15747
Interactions (# of links)	1437	10835	44711	25723	37823	85306
Average connectivity $\langle k \rangle$	3.792	6.387	13.761	6.962	7.3	10.835
Power Law exponent ( $\gamma$ )	1.93	1.88	1.62	1.86	1.79	1.75

Compiled from six databases: DIP, BIND, IntAct, Reactome, HPRD, MINT

enriched into 10 top level molecular functions namely: catalytic activity, structural molecule activity, transporter activity, binding, antioxidant activity, chaperone regulator activity, enzyme regulator activity, transcription regulator activity, translation regulator activity and molecular transducer activity.

#### Studying the protein-partner retention across species

Here the aim was to find if a hub-protein retained the same partner across the species. A statistically significant (e-value less than 0.001) similarity was counted as 1 (i.e., partner retained) else it was counted as 0 (i.e., partner changed). By summing up the scores, the total number of evolutionally conserved interactions was computed. Furthermore, the total number was divided by individual connectivity, and average score was calculated. The statistical significance of ‘getting rich’, ‘getting poor’ and ‘flexible’ groups was studied using F-test. The F-value was calculated as:

$$F = \frac{SS_M/df_M}{SS_G/df_G}$$

$SS_M$  and  $SS_G$  are sum squared error from the total mean and categorical means.  $df_1$  and  $df_2$  are degree of freedoms of  $SS_M$  and  $SS_G$ , respectively, and their values are 20 (21–1) and 2 (3–1) P-value is obtained from F-distribution of respective degrees of freedom (Brandt 1983).

## Result

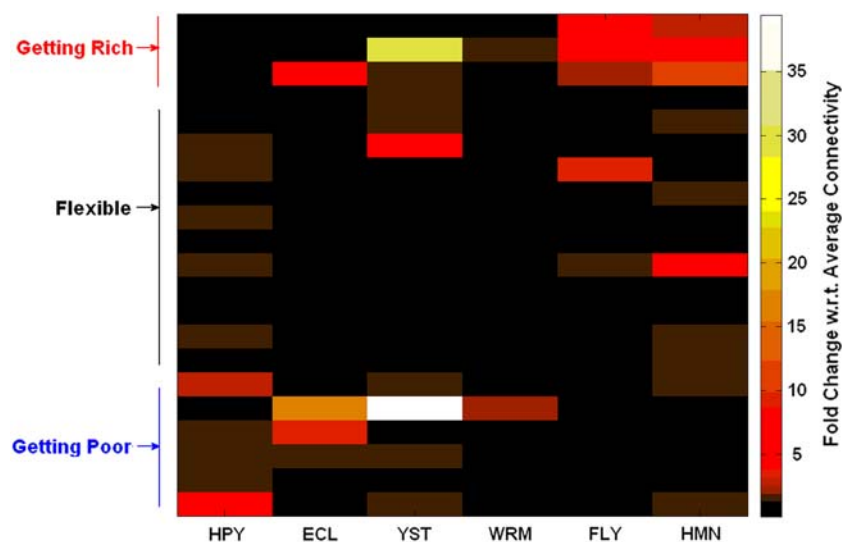
Table 1, summarizes the PPI network data obtained for six model species. All the networks exhibit power law degree distribution of nodes (as  $p(k) \sim k^{-\gamma}$ ) with very similar constant

values for the exponent. This reveals comparable scale free nature of all the PPI networks selected. However, the average connectivity,  $\langle k \rangle$  varies largely across the six species from 3.8 in *H. pylori* (HPY) to 13.8 in *S. cerevisiae*. For ortholog selection, a total of 101 orthologs in Prokaryotic group (denoted as Ortho\_Pk) and 377 in Eukaryotic group (Ortho\_Ek) were identified, out of which 21 were common to both (Ortho\_PkEk). These groups of orthologs form the basis of our analysis and argument in support of hub convertibility.

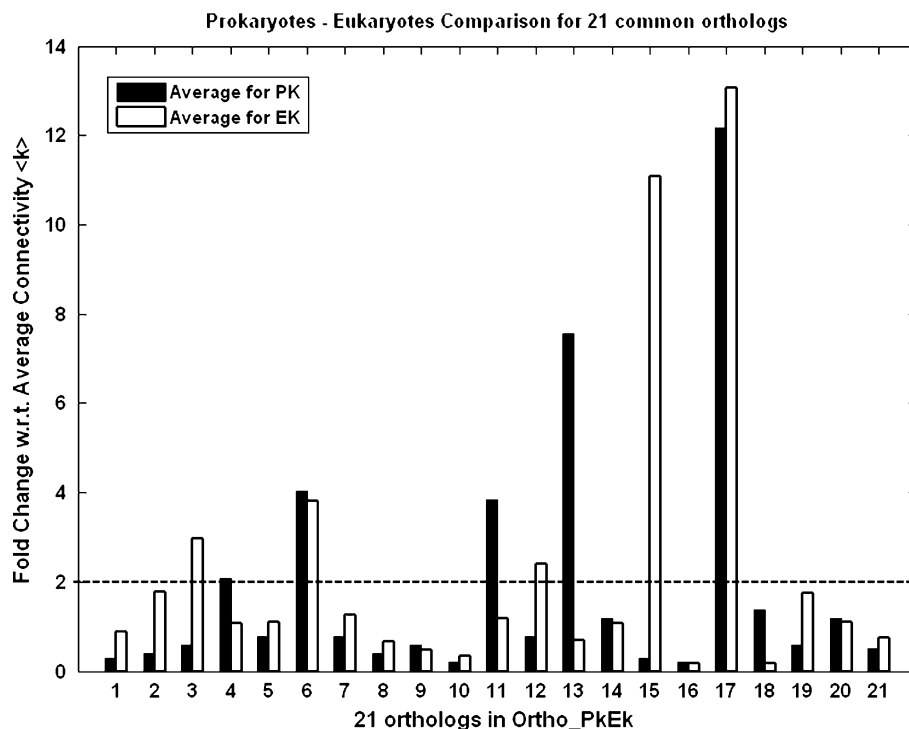
#### Ortho\_PkEk proteins

Figure 1 shows the hub connectivity profile (based on  $k_{fc}$ ) for six species. Using a cutoff level of 2-fold change, all the 21 Ortho\_PkEk proteins showed hubness in at least one species. Furthermore analysis of the kfc data showed three distinct trends for 21 core proteins (separately indicated for clarity in Fig. 1). A total of 57.1% of the core proteins showed “flexible” hub convertibility, 14.3% proteins showed “getting rich” and 28.6% proteins showed “getting poor” trend. It is important to note that “getting poor” proteins show higher confidence of hub convertibility. The hubness is lost in higher organisms even though the cutoff is far below the actual limit ( $k_{fc} \geq 10$  for Eukaryotes). In order to observe the hub conversion between two major groups of phylogeny, we compared the average  $k_{fc}$  ( $\langle k_{fc} \rangle$ ) profiles by grouping species in each category (Fig. 2). A common cutoff of  $\langle k_{fc} \rangle \sim 2$  was used to observe the hub convertibility between Prokaryotes and Eukaryotes. Some proteins [proteins 6 and 17 in Ortho\_EkPk] were found to be retained as highly connected nodes in both the species groups (Hub–Hub). It is evident from their annotation that these two proteins (with functional domains Ribosomal L14 and Elongation Factor, respectively) are essential to

**Fig. 1** Distribution of kfc for 21 Ortho\_PkEk proteins over six species. The profiles are clustered separately to highlight the three distinct trends in connectivity. Groups of 3, 6 and 12 proteins, respectively, show “getting rich”, “getting poor” and “flexible” hub convertibility



**Fig. 2** Distribution of averaged  $k_{fc}$  for 21 Ortho\_PkEk proteins.  $\langle k_{fc} \rangle$  for Prokaryotes (PK) is obtained by averaging  $k_{fc}$  of HPY, ECL and for Eukaryotes (EK) by averaging over YST, WRM, FLY, HMN



protein folding mechanism. Orthologs (12, 15 and 19 with functional annotations DNAj, ATP\_synthase and dehydrogenase) were observed to “get rich”, in Eukaryotes. Orthologs (4, 11 13 representing MSH family, helicase and ClpA protein, respectively) show “getting poor” trend from Prokaryotes to Eukaryotes.

#### Hub convertibility in prokaryotes and eukaryotes

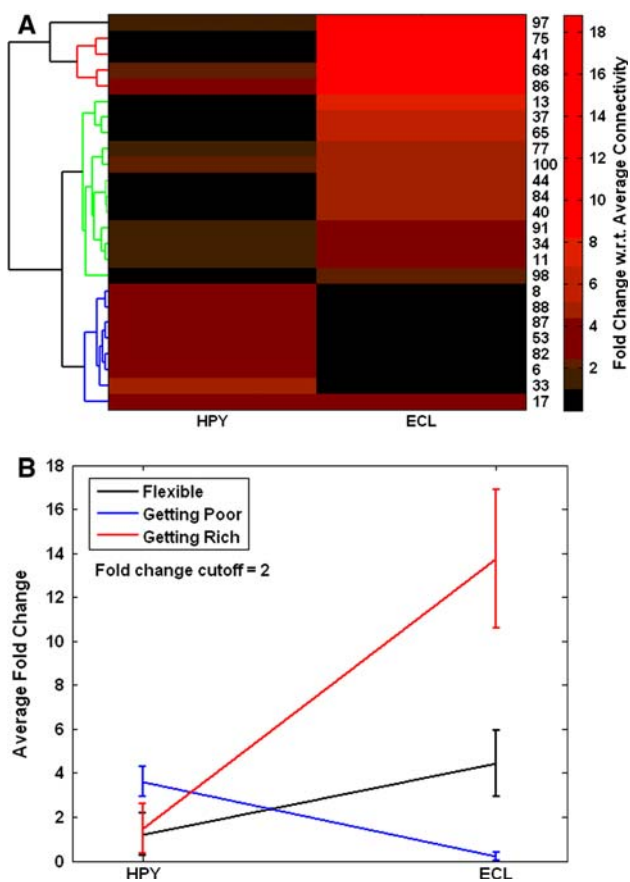
Using the criteria of fold change cutoff  $k_{fc}$  (2 and 10 for Prokaryotes and Eukaryotes respectively), about 25% of Ortho\_Pk proteins (25 out of 101) and 12.7% of Ortho\_Ek proteins (48 out of 377) were found to exhibit hubness. It was difficult to establish meaningful conclusions for the remaining orthologs exhibiting significant connectivity variations. These nodes were below the connectivity threshold set by hubness criteria for all the organisms.

Even though the two organisms (*H. pylori* and *E. coli*) belong to the same super kingdom of Eubacteria, they exhibit different patterns of hubness for the conserved proteins (Fig. 3a). Thirteen out of 25 potential hub candidates from Ortho\_Pk list, exhibit differential hubness. Eight core proteins (32%) are significantly richer (seven times higher) in HPY compared to ECL showing “getting poor” trend. Whereas five proteins (20%) get richer in ECL connectivity (10 times more than HPY  $k_{fc}$  values), the remaining 12 proteins continue to exist as hub nodes in both the networks. Figure 3b, with average  $k_{fc}$  profiles for these three groups summarizes these observations for Prokaryote hub convertibility.

Among 48 Ortho\_Ek hub proteins, 45.8% proteins (22 out of 48) were found to show “get rich” pattern whereas 39.6% of ortholog proteins (19 out of 48) showed decreasing hubness across the four species (Fig. 4a). The gain and loss of hubness indicates a significant change of connectivity in several organisms, with average  $k_{fc}$  profiles (in Fig. 4b) indicating an order of magnitude difference. Hubness trends for the “flexible” nodes shows a large deviation in connectivity, possibly due to presence of higher number of false positives in Yeast and Human data. The “getting poor” and “getting rich” hub convertibility trends are stable with smaller standard deviations. The hub connectivity profile shows a distinct trend among various organisms (Fig. 5a–d).

Figure 6 shows percent protein-partners with at least one significantly similar (E-value of less than 0.001) partner protein. The mean values and standard deviations for each protein class are shown. The figure suggests that getting rich and getting poor categories of proteins are significantly higher in retaining protein-partners than the proteins in the flexible category.

The relative abundance of molecular functions for six species in each category was also studied (Fig. 7a). Figure 7b shows changes in average functional counts from *H. Pylori* to Human for ‘getting rich’, ‘flexible’ and ‘getting poor’ categories. The changes reflect number of documented annotations to a protein. Our work broadly suggests some functional meaning of the protein-partnership trend. However, more studies need to be performed to address this issue in depth.

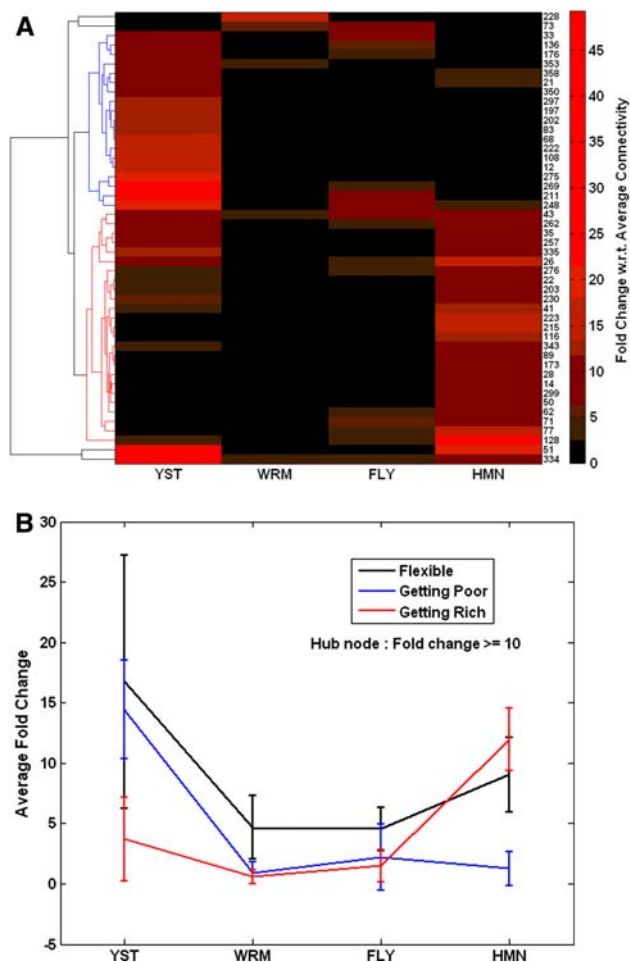


**Fig. 3** **a** kfc distribution of 25 potential hub Ortho\_Pk proteins across *H. pylori* and *E. coli*. The trends are clustered into three distinct groups. Total of 12 proteins show smaller variations (clustered in the middle as green group on the dendrogram), 8 proteins are “getting poor” (blue cluster) and only 5 proteins are “getting rich” (top block in red cluster). The labels on right hand side indicate protein indices in Ortho\_Pk set. **b** Average kfc profiles ( $\pm\sigma$ ) showing the hub convertibility phenomena in Prokaryotes

**Discussion**

The ability to understand information-flow in bio-molecular networks is one of the key goals in systems biology. Hubs are central to this process of information management, as they literally ‘hold the networks together’.

Originally introduced by Barabasi and Albert (1999) and followed by several interesting papers (Barabasi and Oltvai 2004; Nacher and Akutsu 2007), scale freeness has been widely accepted as a generic model of networks exhibiting power law distributions. However, there are some reports of protein interaction networks not conforming to the power law (Khanin and Wit 2006; Tanaka et al. 2005). It is further argued that some published PPI networks are better described by an exponential function and proponents of this approach recommend using rank plots instead of frequency-degree plots (Tanaka et al. 2005). Furthermore, it has been demonstrated that power law degree distribution



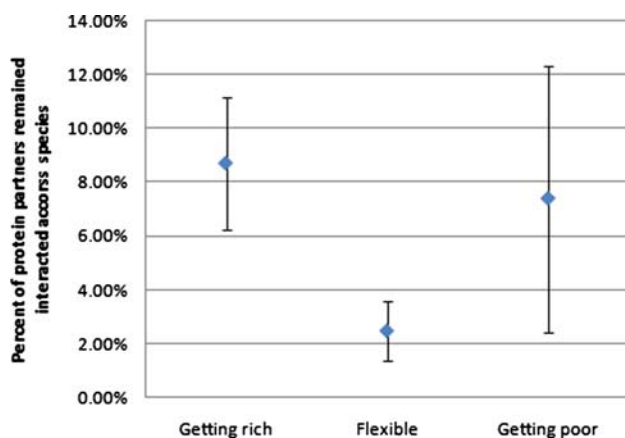
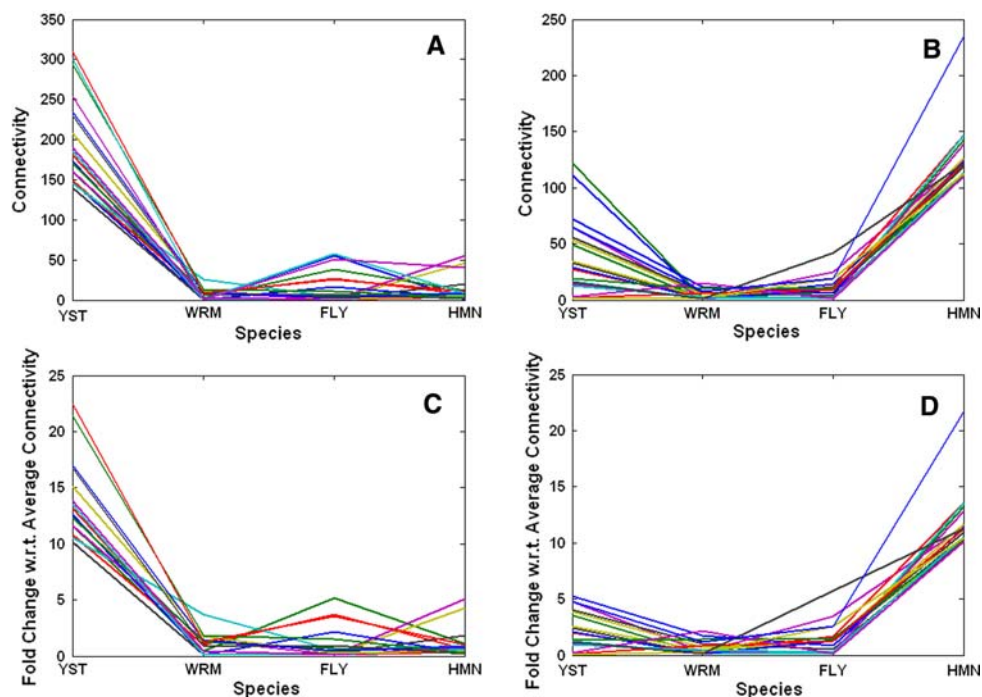
**Fig. 4** **a** kfc distribution of 48 potential hub Ortho\_Ek proteins across four Eukaryote species. Proteins clustered in red group on the dendrogram are “getting rich”, proteins in blue cluster are “getting poor” and others are “flexible”. The labels on right hand side indicate protein indices in Ortho\_Ek set. **b** Average kfc profiles ( $\pm\sigma$ ) showing the hub convertibility phenomena in Eukaryotes

is equivalent to a power law degree-rank function only if scaling exponent is greater than 2 (Wu et al. 2008). It is important to recognize that rich getting richer paradigm is not the only mechanism leading to scale free networks (Li et al. 2005). Finally, the scale-free topology of existing protein-protein interaction networks may not be confidently extrapolated to complete interactomes (Han et al. 2005).

We addressed the paradigm of ‘rich-getting-richer’ from a different perspective. Our aim was to see if rich always get richer. If no, what would happen if hub proteins lost most of the links? As a first step, the data were integrated from six protein-protein interaction databases to create a reasonably large size and variety of the sample. Sequence homology was used to eliminate redundancy among hub proteins. A protein node was classified as a hub or non-hub based on the extent of its connectivity. People have used



**Fig. 5** Connectivity [ $k$ ] profiles for Ortho\_Ek proteins **a** “getting poor”, **b** “getting rich”. Connectivity fold change (kfc) profiles **c** “getting poor”, **d** “getting rich”. Overall increase decrease trends are seen clearer in kfc profiles. kfc cutoff of 10 is used to decide hub nodes



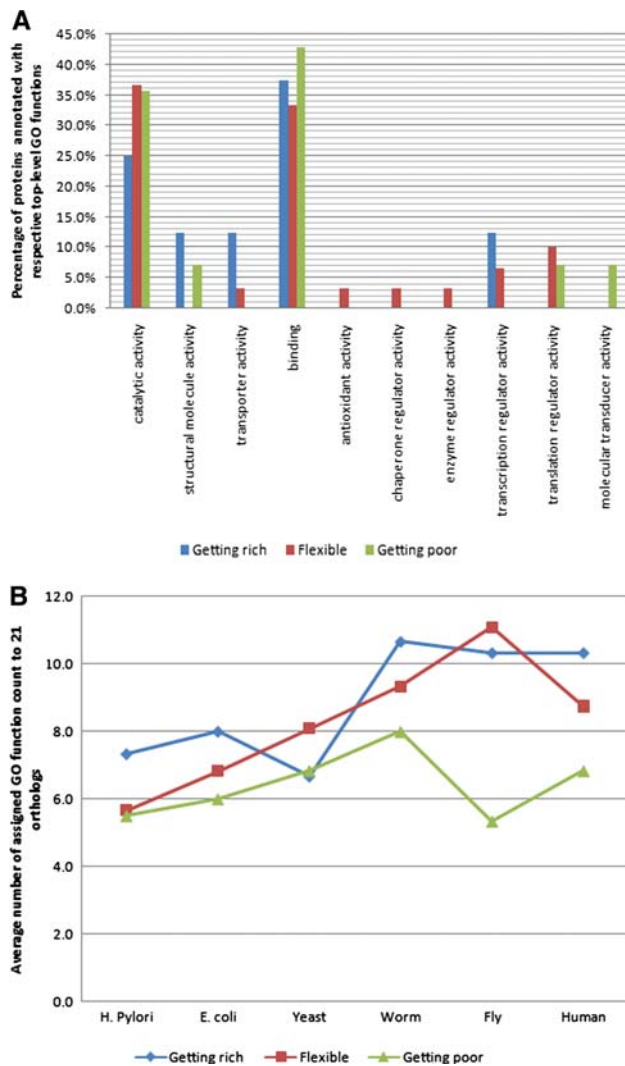
**Fig. 6** The functional classification of Ortho\_Ek core proteins performed using KOG

several criteria to define hubness based on the type of network analysis. For example, Barabasi and Albert (1999) suggest that, hub nodes in scale free networks generally exhibit connectivity,  $k$ , an order of magnitude higher than average vertex degree  $\langle k \rangle$  of the network. Unfortunately, such measures cannot be generalized to biological networks exhibiting modularity. Han et al. (2004) used hub node criterion of  $k \geq 5$  (in a network with average vertex degree  $\langle k \rangle = 3.6$ ). Such arbitrary cutoff measures can lead to misclassification of nodes in large networks with potential false positive interactions. Single criterion, based on  $k$  cutoff is misleading as even a non-hub node in dense network might have edges more than hub node in less

denser network. The  $Z$  score cutoff ( $\geq 2.5$ ), based on the standardized normal distribution of connectivity values ( $k$ ) has also been used to establish significant hub nodes (Ekman et al. 2006; Guimera and Nunes Amaral 2005). The  $Z$  scores definition is inappropriate in our study, since different networks exhibit different degree distributions with varying  $\langle k \rangle$ .

One of the confounding factors in studying PPI networks is the low quality of experimental data (Gentleman and Huber 2007; Jensen and Bork 2008). This can be particularly worrying if one relies too much on a particular database or attempts to integrate several independently constructed databases (Alexeyenko and Sonnhammer 2009). To address the issue of data integrity, we adopted a stringent metric of 10-fold change to identify hubs and minimize false positives, mainly for eukaryotes. Thus, in theory even if 50% edges turn out to be false positives, the protein will still show a significantly high vertex degree to qualify for the standard definition of hubness. A smaller cutoff was, however, used for prokaryotic hub proteins given the relatively smaller  $\langle k \rangle$  and smaller size of networks.

The current dataset includes proteins representing both physical and functional interactions to reduce data bias (Han et al. 2004). To ensure that we were studying the same protein in different organisms, a set of conserved proteins (orthologs) were extracted in all the six species. A stringent threshold of 50 bits and matching segment exceeding 50% of its total length was adopted. The ortholog data were also examined using COG and KOG protein family classification.



**Fig. 7** **a** Functional assignment of 21 core ortholog proteins for the three categories. GO functional assignments are enriched in the top-level hierarchy and averaged by the number of participant proteins for group wise comparison. **b** Quantitative changes in functional assignment of 21 core ortholog protein in all the three categories

Our observation of proteins “getting rich” supports the preferential attachment model for all scale free networks (Barabasi and Albert 1999; Qian et al. 2001). However, we also found incidences of rich-proteins-getting-poor. Interestingly, the existing network growth models do not consider the possibility of fluctuating connectivity in hub proteins. Even if we take into account the yeast, *S. cerevisiae* where average connectivity score is higher, the results are still significant at 10-fold change cutoff. Our findings support the concept of hub convertibility i.e., loss, retention and gain of hubness based on the context.

We further asked if the getting-rich or getting-poor hub proteins retain their core protein partners across all the species? A total of 380 protein-partners were found to exhibit significant similarity by way of their sequence

identity (e-value  $<0.001$ ). Among them, 24 partners shared more than two species. Although statically non-significant ( $P$ -value = 0.13), the getting-rich proteins tend to maintain their interacting partners, thereby reflecting “an intrinsic scale-free design” of larger networks. However from this data, we could not identify an interaction partner that was ortholog in all six species. Interestingly, the functional-spread of the hub proteins shows a decrease as their partners gradually downsize in number. The present hub convertibility data suggests that several ancient proteins (hub nodes in Prokaryotes) are conserved i.e., remain as orthologs in eukaryotes despite the loss of hubness. One reason for this observation could be their key roles (Kunin et al. 2004) in the cell, irrespective of their observed connectivity patterns.

In future, it would be interesting to address following questions (i) how do new hubs arise in the networks? (ii) Which cellular decisions determine the ‘retirement’ of proteins that previously existed as hubs? (iii) Is there any protein-structure basis for hub convertibility? (iv) How networks compensate for the loss of rich proteins getting poor? (v) How is robustness maintained in view of fluctuating connectivity trends in hub/non-hub proteins? (vi) Are some of the proteins more susceptible to hub convertibility than others, If yes—why? (vii) Can the ‘conversion potential’ of a protein be predicted from its sequence/structure data? (viii) Does compartmentalization impact the gain or loss of hubness? (ix) Does protein-partner loss/retention impact the maintenance of specific functional modules? (x) How do metabolic needs impact component reuse vis-à-vis retention or invention of novel functional modules in a network?

**Acknowledgments** This work was funded with intramural grant from RIKEN. P. K. Dhar gratefully acknowledges Dr. Y. Sakaki’s encouragement for this work. R. K. Rao and L. Samavedham thankfully acknowledge the internship support from RIKEN and NUS.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Alexeyenko A, Sonnhammer EL (2009) Global networks of functional coupling in eukaryotes from comprehensive data integration. *Genome Res* 25 (Epub ahead of print)
- Alexeyenko A, Tamas I, Liu G, Sonnhammer EL (2006) Automatic clustering of orthologs and inparalogs shared by multiple proteomes. *Bioinformatics* 22:e9–e15. doi:10.1093/bioinformatics/bt1213
- Bader GD, Betel D, Hogue CW (2003) BIND: the biomolecular interaction network database. *Nucleic Acids Res* 31:248–250. doi:10.1093/nar/gkg056

- Barabasi AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512. doi:[10.1126/science.286.5439.509](https://doi.org/10.1126/science.286.5439.509)
- Barabasi AL, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5:101–113. doi:[10.1038/nrg1272](https://doi.org/10.1038/nrg1272)
- Brandt S (1983) *Statistical and computational methods in data analysis*. North Holland
- Chatr-aryamontri A, Ceol A, Palazzi LM et al (2007) MINT: the Molecular INTeraction database. *Nucleic Acids Res* 35:D572–D574. doi:[10.1093/nar/gk1950](https://doi.org/10.1093/nar/gk1950)
- D'Souza RM, Borgs C, Chayes JT et al (2007) Emergence of tempered preferential attachment from optimization. *Proc Natl Acad Sci USA* 104:6112–6117. doi:[10.1073/pnas.0606779104](https://doi.org/10.1073/pnas.0606779104)
- Eisenberg E, Levanon EY (2003) Preferential attachment in the protein network evolution. *Phys Rev Lett* 91(3):138701.1–4
- Ekman D, Light S, Bjorklund AK, Elofsson A (2006) What properties characterize the hub proteins of the protein–protein interaction network of *Saccharomyces cerevisiae*? *Genome Biol* 7:R45. doi:[10.1186/gb-2006-7-6-r45](https://doi.org/10.1186/gb-2006-7-6-r45)
- Gentleman R, Huber W (2007) Making the most of high-throughput protein–interaction data. *Genome Biol* 8:112. doi:[10.1186/gb-2007-8-10-112](https://doi.org/10.1186/gb-2007-8-10-112)
- Guimera R, Nunes Amaral LA (2005) Functional cartography of complex metabolic networks. *Nature* 433:895–900. doi:[10.1038/nature03288](https://doi.org/10.1038/nature03288)
- Han JD, Bertin N, Hao T et al (2004) Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature* 430:88–93. doi:[10.1038/nature02555](https://doi.org/10.1038/nature02555)
- Han JD, Dupuy D, Bertin N, Cusick ME, Vidal M (2005) Effect of sampling on topology predictions of protein–protein interaction networks. *Nat Biotechnol* 23(7):839–844. doi:[10.1038/nbt1116](https://doi.org/10.1038/nbt1116)
- Haynes C, Oldfield CJ, Ji F et al (2006) Intrinsic disorder is a common feature of hub proteins from four eukaryotic interactomes. *PLOS Comput Biol* 2(8):e100. doi:[10.1371/journal.pcbi.0020100](https://doi.org/10.1371/journal.pcbi.0020100)
- Hermjakob H, Montecchi-Palazzi L, Lewington C et al (2004) IntAct: an open source molecular interaction database. *Nucleic Acids Res* 32:D452–D455. doi:[10.1093/nar/gkh052](https://doi.org/10.1093/nar/gkh052)
- Jensen LJ, Bork P (2008) *Biochemistry*. Not comparable, but complementary. *Science* 322:56–57. doi:[10.1126/science.1164801](https://doi.org/10.1126/science.1164801)
- Joshi-Tope G, Gillespie M, Vastrik I et al (2005) Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res* 33:D428–D432. doi:[10.1093/nar/gki072](https://doi.org/10.1093/nar/gki072)
- Khanin R, Wit E (2006) How scale-free are biological networks. *J Comput Biol* 13(3):810–818. doi:[10.1089/cmb.2006.13.810](https://doi.org/10.1089/cmb.2006.13.810)
- Kunin V, Pereira-Leal JB, Ouzounis CA (2004) Functional evolution of the yeast protein interaction network. *Mol Biol Evol* 21:1171–1176. doi:[10.1093/molbev/msh085](https://doi.org/10.1093/molbev/msh085)
- Li L, Alderson D, Doyle JC, Willinger W (2005) Towards a theory of scale-free graphs: definition, properties, and implications. *Internet Math* 2(4):431–523
- Light S, Kraulis P, Elofsson A (2005) Preferential attachment in the evolution of metabolic networks. *BMC Genomics* 6:159. doi:[10.1186/1471-2164-6-159](https://doi.org/10.1186/1471-2164-6-159)
- Mishra GR et al (2006) Human protein reference database—2006 update. *Nucleic Acids Res* 34:D411–D414. doi:[10.1093/nar/gkj141](https://doi.org/10.1093/nar/gkj141)
- Nacher JC, Akutsu T (2007) Recent progress on the analysis of power-law features in complex cellular networks. *Cell Biochem Biophys* 2007:37–47. doi:[10.1007/s12013-007-0040-7](https://doi.org/10.1007/s12013-007-0040-7)
- O'Brien KP, Remm M, Sonnhammer EL (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res* 33:D476–D480. doi:[10.1093/nar/gki107](https://doi.org/10.1093/nar/gki107)
- Pagel M, Meade A, Scott D (2007) Assembly rules for protein networks derived from phylogenetic–statistical analysis of whole genomes. *BMC Evol Biol* 7(Suppl 1):S16. doi:[10.1186/1471-2148-7-S1-S16](https://doi.org/10.1186/1471-2148-7-S1-S16)
- Qian J, Luscombe NM, Gerstein M (2001) Protein family and fold occurrence in genomes: power-law behaviour and evolutionary model. *J Mol Biol* 31:673–681. doi:[10.1006/jmbi.2001.5079](https://doi.org/10.1006/jmbi.2001.5079)
- Rodriguez-Caso C, Medina MA, Solé RV (2005) Topology, tinkering and evolution of the human transcription factor network. *FEBS J* 272:6423–6434. doi:[10.1111/j.1742-4658.2005.05041.x](https://doi.org/10.1111/j.1742-4658.2005.05041.x)
- Salwinski L, Miller CS, Smith AJ et al (2004) The database of interacting proteins: 2004 update. *Nucleic Acids Res* 32:D449–D451. doi:[10.1093/nar/gkh086](https://doi.org/10.1093/nar/gkh086)
- Singh GP, Ganapathi M, Dash D (2007) Role of intrinsic disorder in transient interactions of hub proteins. *Proteins* 66:761–765. doi:[10.1002/prot.21281](https://doi.org/10.1002/prot.21281)
- Tanaka R, Yi TM, Doyle J (2005) Some protein interaction data do not exhibit power law statistics. *FEBS Lett* 579:5140–5144. doi:[10.1016/j.febslet.2005.08.024](https://doi.org/10.1016/j.febslet.2005.08.024)
- Wu J, Tan Y, Deng H, Zhu D (2008) Relationship between degree–rank function and degree distribution of protein–protein interaction networks. *Comput Biol Chem* 32:1–4. doi:[10.1016/j.compbiolchem.2007.08.008](https://doi.org/10.1016/j.compbiolchem.2007.08.008)