Article

The polarity sub-network in the yeast network of protein-protein interactions

Luca Paris, Gianfranco Bazzoni
Mario Negri Institute of Pharmacological Research, Milano, Italy
E-mail: gianfranco.bazzoni@marionegri.it

Received 3 September 2011; *Accepted 12 October 2011*; *Published online 20 November 2011*
IAEES

Abstract
Rare, but highly connected, hub proteins subdivide hierarchically global networks of interacting proteins into modular clusters. Most biological research, however, focuses on functionally defined sub-networks. Thus, it is important to know whether the sub-networks retain the same topology of the global networks, from which they derive. To address this issue, we have analyzed the protein-protein interaction sub-network that participates in the polarized growth of the budding yeast *Saccharomyces cerevisiae* and that is derived from the global network of this model organism. We have observed that, in contrast to global networks, the distribution of connectivity k (i.e., the number of interactions per protein) does not follow a power law, but decays exponentially, which reflects the local absence of hub proteins. Nonetheless, far from being randomly organized, the polarity sub-network can be subdivided into functional modules. In addition, most non-hub connector proteins, besides ensuring communications among modules, are linked mutually and contribute to the formation of the polarisome, a structure that coordinates actin assembly with polarized growth. These findings imply that identifying critical proteins within sub-networks (e.g., for the aim of targeted therapy) requires searching not only for hubs but also for key non-hub connectors, which might remain otherwise unnoticed due to their relatively low connectivity.

Keywords protein interactions; protein interaction networks; connectivity; topological role; budding yeast; cell polarity.

1 Introduction

Previous studies in the biological, social and information sciences have analyzed the large-scale topology of complex systems, including global networks of protein-protein interactions. In these networks, hub proteins, which are rare proteins with an exceedingly large number of interactions, hold together all the other proteins, which have just few interactions (Barabasi and Oltvai, 2004; Martínez-Antonio, 2011; Tacutu et al., 2011). Accordingly, the connectivity k is not distributed uniformly, but decays as a power law (Barabasi and Albert, 1999; Zhang, 2011). Moreover, hubs organize the global networks hierarchically into functional modules of densely connected proteins (Ravasz et al., 2002). However, it is unclear whether the same topological organization is retained in the functionally and locally defined sub-networks that derive from the global networks.

The issue is important, because biological investigation usually deals with smaller-scale sub-networks (Hartwell et al., 1999), i.e., fractions of global networks that are restricted locally (to cell components) and/or

functionally (to biological processes). Thus, we asked whether the same topological features are retained, when switching from the whole (the network) down to the parts (the sub-networks).

To answer this question, we have analyzed a sub-network of protein-protein interactions, which is restricted locally (to the growing bud) and functionally (to polarized growth), is well defined (in the extensively studied model organism *Saccharomyces cerevisiae*) (Drubin and Nelson, 1996; Bork et al., 2004) and is subdivided into functional modules (Drees et al., 2001; Gavin et al., 2006). We have identified a hub-less and non-hierarchical organization of the polarity sub-network into highly inter-connected modules.

2 Methods

2.1 Generation of the polarity and random networks

To generate the polarity network, protein-protein interactions (supported by a published reference) were obtained by searching reviews (Pruyne and Bretscher, 2000a, b), articles (Drees et al., 2001; Gavin et al., 2006; Collins et al., 2007) and the *Saccharomyces Genome Database* (Cherry et al., 1997). To generate the random network, two series of 117 numbers n each ($1 \leq n \leq 117$) were distributed in a uniform manner, with the numbers in each pair of the two series corresponding to pairs of interacting nodes.

2.2 Network fragmentation

Network fragmentation was performed as described (Albert et al., 2000). Briefly, nodes were removed sequentially and, at each step t, the number n of nodes remaining in the fragments was counted. The size S of the largest fragment was expressed as the ratio between $n(t)$ and the initial number of nodes $n(t_0)$. Finally, S was plotted as function of the fraction f of removed nodes.

2.3 Clustering analysis and topological roles

Clustering was performed as described (Rives and Galitski, 2003). Briefly, networks were represented as undirected graphs of nodes linked by edges, and each edge was assigned a length of 1. Then, an adjacency matrix of all-pairs shortest path distances was subjected to average linkage hierarchical clustering (Eisen et al., 1998) with uncentered correlation as the similarity metric, using the Cluster 3.0 (de Hoon et al, 2004) and Java Tree View 1.0.13 (Saldanha, 2004) programs for hierarchical clustering and visual display of the clustered tree, respectively. Cytoscape 2.6.1 (Shannon et al., 2003) was used to visualize the networks. Finally, the topological role of each node i was defined by measuring its within-module degree z_i and participation coefficient P_i, as described (Guimera and Nunes Amaral, 2005).

3 Results and Discussion

3.1 Exponential distribution of connectivity in the polarity network

Starting from the global network of protein-protein interactions in *S. cerevisiae*, a list was made of all the interactions among the proteins that contribute to polarized growth in the yeast bud (*Online material 1*). With this list, a sub-network was assembled, which hereafter is referred to as 'polarity network'. The polarity network, which consists of N nodes (117 proteins) and E edges (285 interactions), has average connectivity ($\langle k \rangle = 2E/N$) of 4.9 edges per node (Fig. 1a *inset*). For comparison, a random network, with the same $\langle k \rangle$ as the polarity network, was generated by uniformly distributing 285 edges among 117 nodes (Fig. 1b *left inset*). In global networks, the degree distribution $P(k)$, which is the probability that a selected protein has exactly k interactions, decays as a power law over several orders of magnitude (Barabasi and Albert, 1999). However, when plotting $P(k)$ as a function of k, we observed that both $P(k)_{\text{polarity}}$ (Fig. 1a) and $P(k)_{\text{random}}$ (Fig. 1b) decay exponentially, which is consistent with exponential (rather than power law) distribution. In particular, the $P(k)_{\text{polarity}}$ curve is fitted by the exponential function

$$P(k)_{\text{polarity}} = \mathrm{P}_0 \, e^{-\mathrm{A}k}$$

(P_0 = 0.233±0.020; A = 0.132±0.046), with an R-squared value of 0.932. In contrast, when fitting the curve by a power law function, the R-squared value is 0.708. As the probability of finding a node with exceedingly high connectivity is practically prohibited in exponential networks, it follows that hub proteins are absent from the polarity network. Remarkably, exponential distribution also characterizes the random network shown in Fig. 1b, as well as other random graphs (*e.g.* the Erdös-Rényi model) that follow a Poisson distribution of connectivity peaked at $<k>$ (Erdös and Rényi, 1960). This observation, however, does not necessarily imply that the polarity network is a random graph. Actually, when fitting the $P(k)_{polarity}$ to a Poisson distribution, we obtained a chi-squared value (X^2_{obs}) of 115.6 and $P(X^2 \geq 115.6) < 10^{-20}$. In contrast, when fitting the $P(k)_{random}$ to a Poisson distribution, we obtained a X^2_{obs} value of 1.2 and $P(X^2 \geq 1.2) = 0.99$, thus indicating that the random (but not the polarity) network follows a Poisson distribution.

Notably, as the interactions used to assemble the polarity network are derived from high-quality data (such as published studies and manually annotated databases), it is unlikely that our analysis is significantly affected by possible incompleteness (false negatives) and incorrectness (false positives) of the initial data. Actually, error analysis confirmed our conclusions about the exponential and non-random nature of the polarity network (*Online material 2*).

3.2 Removal of the most connected proteins does not cause massive fragmentation of the polarity network

Even though proper hub proteins (with $k >> <k>$) are absent from the polarity network, the exponential tail of $P(k)_{polarity}$ indicates the presence of few nodes with $k > <k>$. It is known that the hubs play a major role in ensuring physical cohesion of the scale-free networks, because these networks become fragmented upon removal of even small fractions of the hubs (Albert et al., 2000). Thus, to test whether the highly connected (albeit non-hub) proteins might play in the polarity network the same role that the hubs play in scale-free networks, we performed a targeted attack, by sequentially removing the most connected proteins in decreasing order of connectivity *k*. As control, we performed a non-targeted attack, which consists of removing the same number of nodes as in the targeted attack, but irrespective of the connectivity. We found that targeted removal of the highly connected proteins did not cause massive fragmentation of the polarity network (*Online material 3*). Specifically, up to about 10% of removed nodes, both targeted and non-targeted attacks caused comparable fragmentation of the polarity and random networks. Only with the removal of higher fractions of nodes, fragmentation became more obvious in the polarity network under targeted (compared with non-targeted) attack, as well as in the polarity (compared with the random) network under targeted attack.

3.3 The polarity network can be subdivided into protein clusters that correspond to functional modules

Another role of the hubs (at least in hierarchical systems) consists of organizing networks into clusters of nodes (Ravasz et al., 2002). A parameter that quantifies the tendency of each node to form clusters is the clustering coefficient *C*. Specifically, if node *a* is directly connected to nodes *b* and *c*, then C_a is the probability that *b* and *c* are also directly connected to each other. Compared with the majority of nodes, the hubs of hierarchical scale-free networks display not only higher *k*, but also lower *C*. As a consequence, $C(k)$, *i.e.* the distribution of *C* as a function of *k*, decreases with *k* and approximates k^{-1} (Watts and Strogatz, 1998). In contrast, in the polarity network, we observed that $C(k)$ does not change with *k* (Fig. 1c).

Taken together, the exponential decay of $P(k)$ and the independence of $C(k)$ on *k* indicate the absence of highly-connected and poorly-clustered hubs. On one side, these observations might suggest that the polarity network does not display the hierarchical structure of many scale-free systems. On the other side, the average clustering coefficient $<C>$, which characterizes the overall tendency of all nodes to form clusters of nodes, is about 10-fold greater in the polarity (0.208±0.022) than in the random (0.027±0.004) network. Thus, in

contrast to scale-free networks (Ravasz et al, 2002), the global lack of hierarchy in the polarity network does not negate its local organization into clusters.

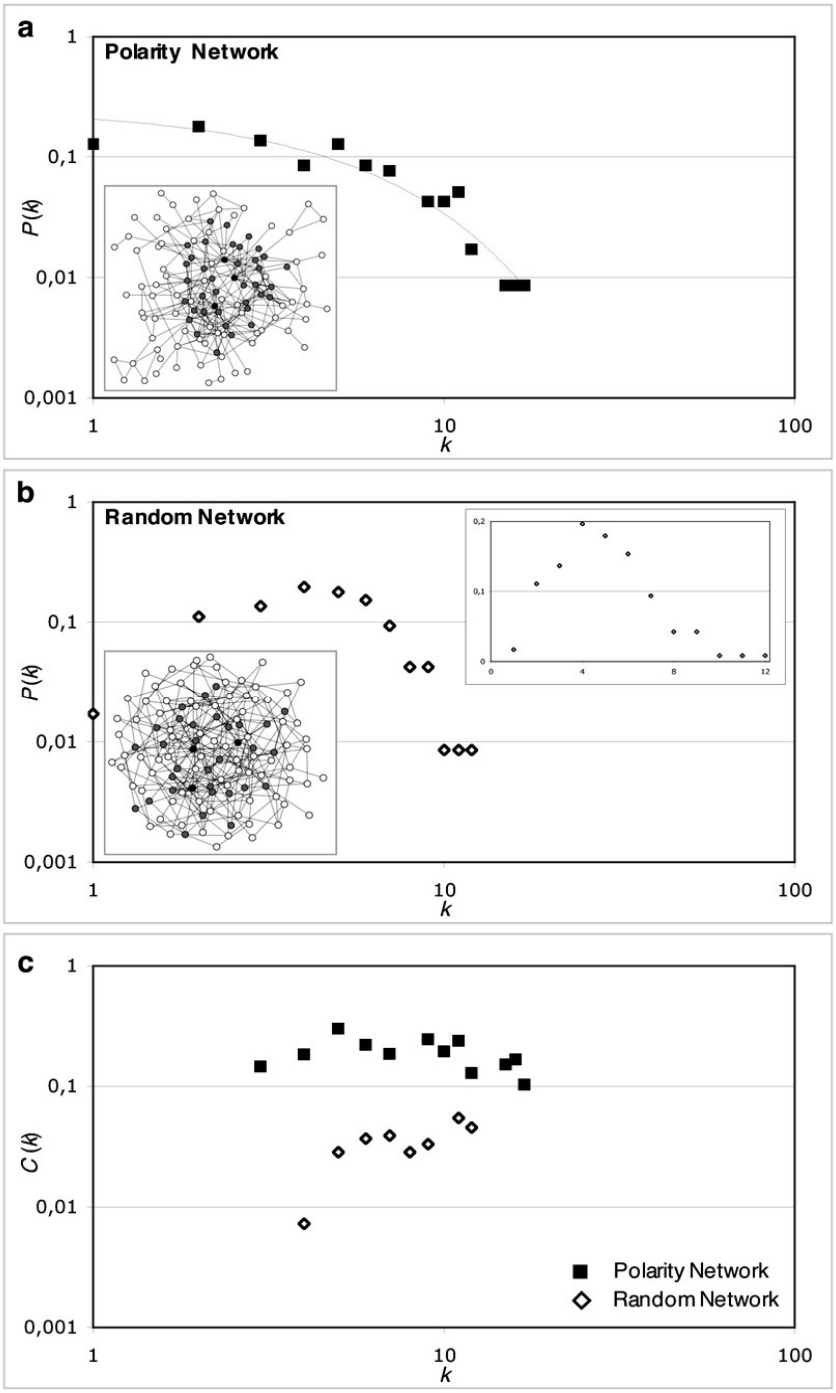

Fig. 1 Distribution of the connectivity and the clustering coefficient. The *P*(*k*) of the polarity (a) and random (b) network is shown in logarithmic plots and, for the random network, also on a linear plot, to highlight that its distribution is peaked at <*k*> (*right inset* of b). Both networks are also displayed visually (*left insets* of a and b), to show that comparable fractions of first neighbours (*gray*) are linked to the most connected nodes (*black*) in the polarity (30,8%) and random (26,5%) network. The independence of *C*(*k*) on *k* in both networks is shown in (c).

Clusters of interconnected nodes that perform distinct functions are often referred to as modules (Hartwell et al., 1999). To examine whether the polarity network can be subdivided into modules, we deployed a clustering method (Rives and Galitski, 2003), which consists of measuring the length *l* of the shortest path (*i.e.* the path with the smallest number of links) between all pairs of nodes in the network. The method is based on the assumption that nodes with similar (*i.e.* clustered) *l* profiles are members of the same module. We found that the polarity network can be subdivided into five clusters (Fig. 2a and *Online material 4*) and that the subdivision is highly modular, as defined by the Girvan and Newman algorithm, which finds the division of a network into modules that maximizes a coefficient *Q* (Newman and Girvan, 2004). The normalized Q_m coefficient of the polarity network (0.57) is comparable with the coefficients of highly modular systems, such as the transcriptional (0.54), neuronal (0.54) and signalling (0.58) networks (Kashtan and Alon, 2005). Furthermore, the clusters correspond to functional modules, as several proteins with a common Gene Ontology (GO) annotation (Ashburner et al., 2000) belong to the same protein cluster (Fig 2b and *Online material 5*).

3.4 Connecting roles of the polarity proteins

To reconcile the modularity of the polarity network with the absence of hubs, we hypothesized the existence of connectors, *i.e.* non-hub nodes that link modules (in spite of a relatively low *k*). To test this hypothesis, we examined the connecting role of each protein based on its pattern of intra- and inter-module connections, using the within-module degree *z* and the participation coefficient *P*, respectively (Guimera and Nunes Amaral, 2005). In this functional cartography, nodes with $z \geq 2.5$ are defined as module hubs and nodes with $z < 2.5$ as non-hubs. In keeping with the above findings, only two proteins (*i.e.* Las17p and Cdc42p) attain a *z* score that is just above the threshold of 2.5 (Fig. 3a). In addition, the coefficient *P* defines the role of each node in connecting to nodes in modules other than its own. In this way, we found that the two module hubs do not play a major connecting role. On one side, Las17p is a provincial hub (*i.e.* a hub with almost all its links within its module; $P \leq 0.30$). On the other side, Cdc42p is certainly a connector hub (*i.e.* a hub with many links to most of the other modules; $0.30 < P \leq 0.75$), but its *P* coefficient is just above the threshold of 0.30. All the other proteins in the polarity network are non-hubs ($z < 2.5$). Among these non-hubs, the vast majority does not play a significant connecting role. More than 90% are either ultra-peripheral or peripheral (*i.e.* nodes with all or most links within their module, respectively; $P \leq 0.05$ or $0,05 < P \leq 0.62$). The remaining non-hub proteins are connectors (*i.e.* nodes with many links to other modules; $0.62 < P \leq 0.80$). At variance with the polarity network, the random network displayed a different pattern (Fig. 3b). Specifically, we found a 4-fold decrease and a 4.4-fold increase in the fractions of ultra-peripheral and connector non-hubs, respectively. Thus, the connecting roles are not randomly distributed among the proteins of the polarity network, thus unveiling the fundamental role of the non-hub connectors in its modular (albeit not hierarchical) organization.

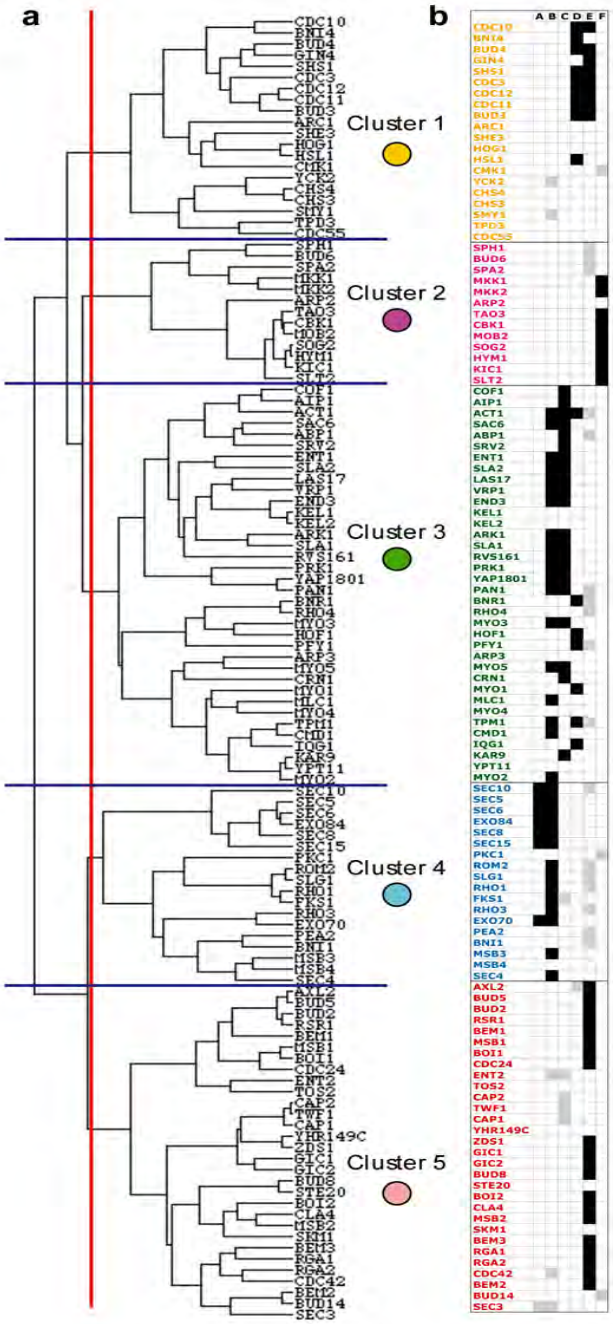

Fig. 2 The polarity network can be subdivided into protein clusters that correspond to functional modules. The network was analyzed by hierarchical clustering to obtain the tree shown in (a). Columns in (b) correspond to six classes of GO terms (A: Exocyst; B: Vesicle-mediated transport; C: Cell cortex, Actin cortical patch; D: Septin ring, Cellular bud neck septin ring, Cellular bud neck contractile ring, E: Establishment of cell polarity, Cellular bud site selection, Axial bud site selection; F: Signal transduction, Cellular morphogenesis). Fillings indicate that the GO terms that annotate a given protein either are (*black*) or are not (*gray*) enriched in the protein cluster, to which the annotated protein belongs. Further details can be found in *Online material 5* and *Fig. S2*.

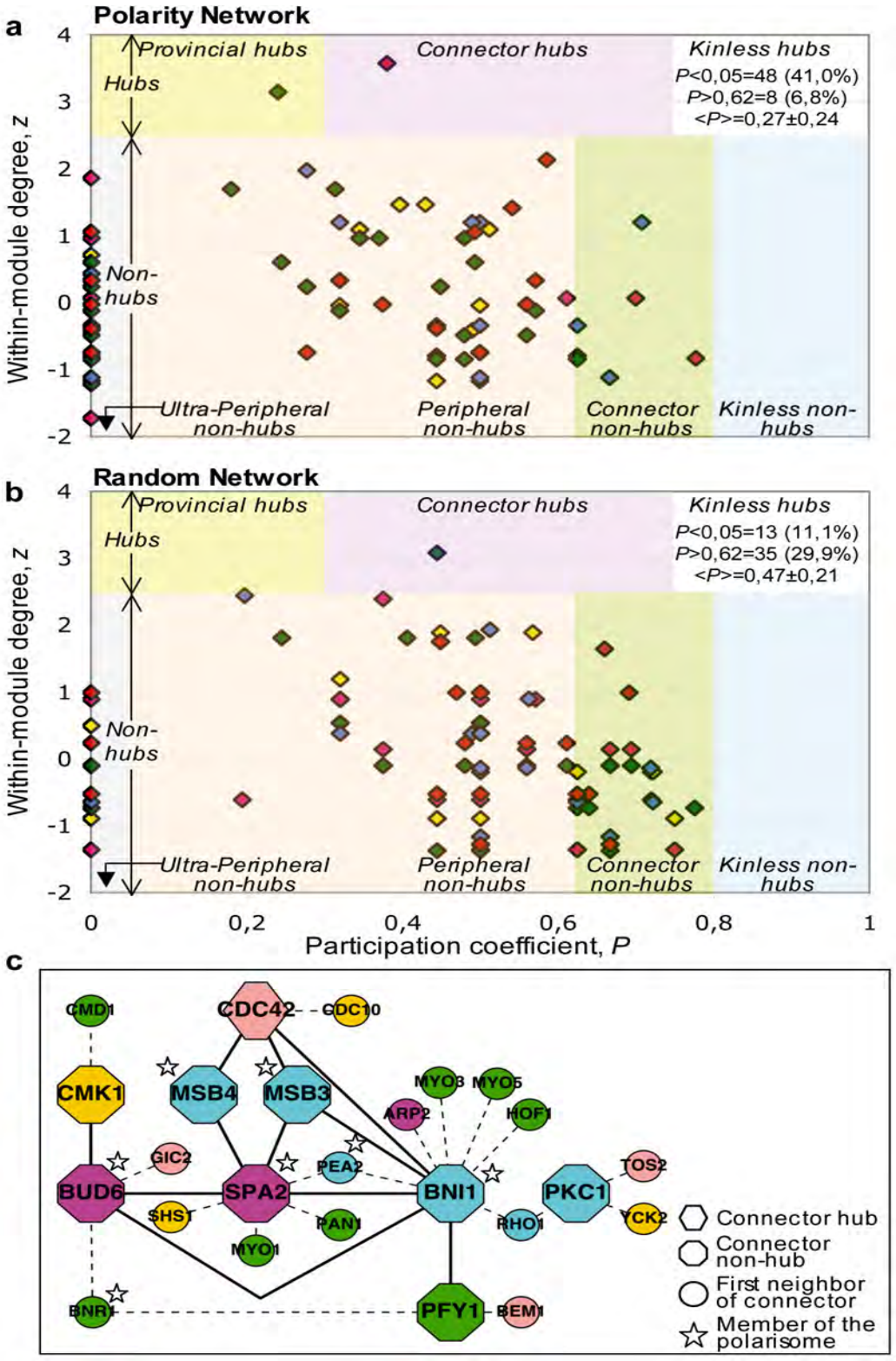

Fig. 3 Connecting roles of the polarity proteins. According to the within-module degree *z* and participation coefficient *P*, individual proteins in the polarity (a) and random (b) network are assigned to seven different roles of connection. (c) A core of connectors links proteins involved in polarized actin assembly. Solid lines indicate direct links between the connectors. Nodes are colour-coded as in Fig 2.

Biological evidence strengthens the topological significance of the identified non-hub connectors. The eight non-hub connector proteins ($z<2.5$; $0.62<P\leq0.80$) are Bni1p, Msb3p, Msb4p and Pkc1p (cluster 1), Cmk1p (cluster 3), Bud6p and Spa2p (cluster 4), and Pfy1p (cluster 5). As the hub Cdc42p (cluster 2) is a connector too, all the five clusters identified above (Fig. 2a) contain at least one connector (Fig. 3c). Noteworthy, eight out of nine connectors are directly associated to each other, thus constituting a central core of connection for the five clusters, while the ninth connector (Pkc1p) is indirectly linked to the core via a common neighbour (Rho1p). In addition, five connectors in the central core (Bni1p, Msb3p, Msb4p, Bud6p and Spa2p), together with two neighbours (Bnr1p and Pea2p), are the components of the polarisome, a structure required for the polarized organization of actin (Pruyne and Bretscher, 2000b; Tcheperegine et al., 2005). In particular, the polarisome provides a docking site for the formin Bni1p, which in turn links the Cdc42p-triggered establishment of polarity with the profilin Pfy1p-mediated polymerization of actin (Evangelista et al., 1997; Sheu et al., 1998). Thus, the connectors do represent a core of functionally important connections between the modules of the polarity network.

3.5 Final remarks

In this study, we have shown that the connectivity distribution within a local sub-network does not follow a power law, but decays exponentially. While this observation reflects the absence of hubs, it does not imply that the polarity network is randomly organized. Rather, we provide evidence that it is orderly subdivided into functional modules, with the non-hub connector proteins taking on the functionally-important role of ensuring inter-module communication (Guimera and Nunes Amaral, 2005). Notably, from all the proteins of the polarity network, this approach has singled out a set of key connectors (mostly members of the polarisome) that link and coordinate different events in the establishment of polarity.

In practice, our findings highlight the usefulness of combining the methods of network clustering (Fig. 2a) (Rives and Galitski, 2003) and role attribution (Fig. 3a) (Guimera and Nunes Amaral, 2005) to identify important proteins within defined sub-networks, based on their ability to connect modules and not merely on their intrinsic connectivity. In a broader perspective, we envision our findings as a new framework for studies of defined sub-networks, such as those associated with biological processes and subcellular organelles (Rachlin et al., 2006; Goemann et al., 2011; Jains et al, 2011; Rodriguez and Infante, 2011), including the junctional complex (Zaidel-Bar et al., 2007; Paris and Bazzoni, 2008; Cirillo and Prime, 2009), as well as human diseases (Vidal et al., 2011; Ibrahim et al, 2011).

Acknowledgements

This work was supported by grants of The Weizmann-Negri Foundation and Association International for Cancer Research Grant (to GB) and a fellowship of Fondazione Musarra (to LP).

References

Albert R, Jeong H, Barabasi AL. 2000. Error and attack tolerance of complex networks. Nature, 406: 378-382

Ashburner M, Ball CA, Blake JA, et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics, 25: 25-29

Barabasi AL, Albert R. 1999. Emergence of scaling in random networks. Science, 286: 509-512

Barabasi AL, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. Nat Rev Genet, 5: 101-113

Bork P, Jensen LJ, von Mering C, et al. 2004. Protein interaction networks from yeast to human. Current Opinion in Structure Biology, 14: 292-299

Cherry JM, Ball C, Weng S, et al. 1997. Genetic and physical maps of Saccharomyces cerevisiae. Nature, 387: 67-73

Cirillo N, Prime SS. 2009. Desmosomal interactome in keratinocytes: a systems biology approach leading to an understanding of the pathogenesis of skin disease. Cellular and Molecular Life Sciences, 66: 3517-3533

Collins SR, Kemmeren P, Zhao XC, et al. 2007. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. Molecular and Cellular Proteomics, 6: 439-450

de Hoon MJ, Imoto S, Nolan J, et al. 2004. Open source clustering software. Bioinformatics, 20: 1453-1454

Drees BL, Sundin B, Brazeau E, et al. 2001. A protein interaction map for cell polarity development. Journal of Cell Biology, 154: 549-571

Drubin DG, Nelson WJ. 1996. Origins of cell polarity. Cell, 84: 335-344

Eisen MB, Spellman PT, Brown PO, et al. 1998. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of U. S. A., 95: 14863-14868

Erdös P, Rényi A. 1960. On the evolution of random graph. Publ Math Inst Hung Acad Sci, 5: 17-61

Evangelista M, Blundell K, Longtine MS, et al. 1997. Bni1p, a yeast formin linking cdc42p and the actin cytoskeleton during polarized morphogenesis. Science, 276: 118-122

Gavin AC, Aloy P, Grandi P, et al. 2006. Proteome survey reveals modularity of the yeast cell machinery. Nature, 440: 631-636

Goemann B, Wingender E, Potapov AP. 2011. Topological peculiarities of mammalian networks with different functionalities: transcription, signal transduction and metabolic networks. Network Biology, 1(3-4): 134-148

Guimera R, Nunes Amaral LA. 2005. Functional cartography of complex metabolic networks. Nature, 433: 895-900

Hartwell LH, Hopfield JJ, Leibler S, et al. 1999. From molecular to modular cell biology. Nature, 402: C47-52

Ibrahim SS, Eldeeb MAR, Rady MAH. 2011. The role of protein interaction domains in the human cancer network. Network Biology, 1(1): 59-71

Jain S, Bhooshan SV, Naik PK. 2011. Mathematical modeling deciphering balance between cell survival and cell death using insulin. Network Biology, 1(1): 46-58

Kashtan N, Alon U. 2005. Spontaneous evolution of modularity and network motifs. Proceedings of the National Academy of Sciences of U. S. A., 102: 13773-13778

Martínez-Antonio A. 2011. *Escherichia coli* transcriptional regulatory network. Network Biology, 1(1): 21-33

Newman ME, Girvan M. 2004. Finding and evaluating community structure in networks. Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 69: 026113

Paris L, Bazzoni G. 2008. The protein interaction network of the epithelial junctional complex: a system-level analysis. Molecular Biology of the Cell, 19: 5409-5421.

Pruyne D, Bretscher A. 2000a. Polarization of cell growth in yeast. Journal of Cell Science, 113(Pt 4): 571-585

Pruyne D, Bretscher A. 2000b. Polarization of cell growth in yeast. I. Establishment and maintenance of polarity states. Journal of Cell Science, 113 (Pt 3): 365-375

Rachlin J, Cohen DD, Cantor C, Kasif S. 2006. Biological context networks: a mosaic view of the interactome. Molecular Systems Biology, 2: 66

Ravasz E, Somera AL, Mongru DA, et al. 2002. Hierarchical organization of modularity in metabolic networks. Science, 297: 1551-1555

Rives AW, Galitski T. 2003. Modular organization of cellular networks. Proceedings of the National Academy of Sciences of U. S. A., 100: 1128-1133

Rodriguez A, Infante D. 2011. Characterization *in silico* of flavonoids biosynthesis in *Theobroma cacao* L.

Network Biology, 2011, 1(1): 34-45

Saldanha AJ. 2004. Java Treeview--extensible visualization of microarray data. Bioinformatics, 20: 3246-3248

Shannon P, Markiel A, Ozier O, et al. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research, 13: 2498-2504

Sheu YJ, Santos B, Fortin N, et al. 1998. Spa2p interacts with cell polarity proteins and signaling components involved in yeast cell morphogenesis. Molecular and Cellular Biology, 18: 4053-4069

Tcheperegine SE, Gao XD, Bi E. 2005. Regulation of cell polarity by interactions of Msb3 and Msb4 with Cdc42 and polarisome components. Molecular and Cellular Biology, 25: 8567-8580

Tacutu R, Budovsky A, Yanai H, et al. 2011. Immunoregulatory network and cancer-associated genes: molecular links and relevance to aging. Network Biology, 1(2): 112-120

Vidal M, Cusick ME, Barabasi AL. 2011. Interactome networks and human disease. Cell, 144: 986-998

Watts DJ, Strogatz SH. 1998. Collective dynamics of 'small-world' networks. Nature, 393: 440-442

Zaidel-Bar R, Itzkovitz S, Ma'ayan A, et al. 2007. Functional atlas of the integrin adhesome. Nature Cell Biology, 9: 858-867

Zhang WJ. 2011. Constructing ecological interaction networks by correlation analysis: hints from community sampling. Network Biology, 1(2): 81-98