Analysis of transcriptional and post-transcriptional regulatory networks

Forcato M[1], Biasiolo M[2], Possamai L[3], Ferrari F[2], Agnelli L[4], Neri A[4], Marchiori M[3], Bortoluzzi S[2], Bicciato S[1]

Motivation

The reconstruction of gene regulatory networks from expression data is a crucial step to unveil the molecular mechanisms governing biological systems. However, the overwhelming number of genes and interactions still hampers the identification of relevant targets and relationships from such large systems. Standard approaches for the analysis of molecular networks aim at identifying targets among the most connected genes (hubs) or at detecting groups of genes organized in sub-networks. In both cases, the goals are the extraction of previously unknown relationships, the identification of regulatory modules involving genes and their common targets, and, ultimately, the formulation of novel mechanistic hypotheses. Although effective, both strategies require the prior knowledge of the genes of interest, thus hampering the capacity of extracting de-novo knowledge from the network. A way to overcome this limitation could be adapting techniques commonly used in the analysis of communication and infrastructure networks. In these fields, a key analysis is the resilience of the network to external disturbances and to malfunctioning. Network robustness strongly relies on the network structure and, in particular, on the existence of paths between the nodes. When nodes or links are removed, the lengths of these paths can increase and some nodes may become disconnected. It is therefore interesting to find the critical components of the network, i.e. nodes or edges that are structurally relevant for the functioning of the network. Applying this concept to gene regulatory networks, critical nodes and edges are critical genes and regulatory interactions, respectively. Usually the most important nodes are considered the most connected ones, i.e. those with the highest degree (hubs). However, in genetic networks, a gene, although not biologically relevant in the analyzed phenotype, may be connected to many other genes simply because is a transcription factor, that normally controls many targets. Here, the analysis of the network critical components is applied to the nodes of transcriptional and post-transcriptional regulatory networks reconstructed from mRNA and miRNA expression data with the aim of identifying genes that are

[1] Department of Biomedical Sciences, Center for Genomic Research, University of Modena and Reggio Emilia, Modena, Italy [2] Department of Biology, University of Padova, Padova, Italy [3] Department of Pure and Applied Maths, Padova, Italy [4] Department of Medical Sciences, Fondazione IRCCS Ospedale Policlinico, Milano, Italy

critical for the structure of the networks, irrespectively of the number of their ingoing or outgoing links.

Methods

Gene (mRNA) and miRNA expression data have been obtained from multiple myeloma (MM) specimens using Affymetrix HG-U133A and Agilent Human miRNA Microarray V2 arrays, respectively. The transcriptional regulatory network has been reconstructed with ARACNe (Basso et al., 2005) whereas the post-transcriptional regulatory network was reconstructed calculating the Pearson correlation coefficient of the expression vectors of miRNA target genes. The analysis of network critical components has been conducted removing each node of the network and calculating the efficiency of the network without that node. The most critical genes have been defined as those nodes whose removal caused the largest drop in network efficiency (Crucitti et al., 2004).

Results

The analysis of critical components revealed that genes with a limited number of connections could be critical for the structure of the network. In particular, the comparison between the node rankings calculated according to node degree and the node criticality in the random, transcriptional, and post-transcriptional networks indicated that i) in a random network there is no significant difference between node degree and node criticality and ii) the vast majority of hub genes are critical nodes, i.e., the nodes whose removal causes a large drop in global efficiency correspond to the most connected genes. Nevertheless, there exists a fraction of nodes that are characterized by a criticality value higher than expected according to their degree. For instance, in the MM transcriptional network, BLNK gene emerged as one of the most critical genes although being connected to only 32 other nodes (as compared to hubs characterized by more than 100 links). Moreover, hubs are not necessarily critical nodes. Indeed, about one half of the most connected nodes in each considered network were not included in the corresponding list of critical nodes. The non-hub critical nodes would have been disregarded as putative regulatory targets due to their limited number of connections although they may provide clues to the detection of key regulatory circuits. Finally, the integration of the transcriptional and post-transcriptional levels allowed identifying critical genes for both types of regulatory interactions and dissecting direct critical relationships at transcriptional level from interaction that are instead indirect, since mediated by post-transcriptional regulation (Biasiolo et al., 2010).

Contact e-mail

silvio.bicciato@unimore.it

Supplementary information

References

[1] Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. Reverse engineering of regulatory networks in human B cells. Nat Genet. 2005 Apr;37(4):382-90 [2] Crucitti P, Latora V, Marchiori M, Rapisarda A. Error and attack tolerance of complex networks. Physica A. 2004;340:388-394 [3] Biasiolo, M, Forcato, M, Possamai, L, Ferrari, F, Agnelli, L, Lionetti, M, Todoerti, K, Neri, A, Marchiori, M, Bortoluzzi, S, Bicciato, S. Critical analysis of transcriptional and post-transcriptional regulatory networks in multiple myeloma. Pac Symp Biocomput. 2010:397-408