Journal of Biometrics & Biostatistics
ISSN: 2155-6180

Gurami et al., J Biom Biostat 2018, 9:2
DOI: 10.4172/2155-6180.1000390

Short Communication Open Access

Replacement of Directed Graph by Acyclic Directed Graph and Its Application in Biostatistics

Tsitsiashvili Gurami[1]*, Bulgakov Victor[2] and Losev Alexandr[1]

[1]Institute for Applied Mathematics FEB RAS, Vladivostok, Russia
[2]Institute of Biology and Soil Science FEB RAS, Vladivostok, Russia

Abstract

In this paper, an algorithm of directed graph replacement by acyclic directed graph is constructing and is applying for a study of the key players required for connecting ABA signaling and ABA-mediated drought and thermo tolerance.

Keywords: Factorization; Cluster; Thermo stability; Acyclic digraph; Protein network

Formulation of Problem

In the analysis of protein networks, it is necessary to develop a procedure for simplifying them by analogy with the method of principal components in mathematical statistics. This simplification can be achieving by removing feedbacks from the network and transformation of the network (directed graph, digraph) into acyclic digraph connecting input and output nodes only without any feedbacks. Possibilities of acyclic digraphs application in calculation of flows intensities are described earlier [1]. In this paper algorithm of the replacement is constructing and is applying for a study of the key players required for connecting ABA signaling and ABA-mediated drought and thermo stability [2-5].

Algorithm of Network Simplification

Replacement of clusters by acyclic digraphs in network

For any digraph, *A* we denote *P(A)* the set of vertices and *Q(A)* the set of edges. Let be G an digraph, produce its clustering with respect to cyclic equivalence using the algorithm of [?]. In each cluster *g* of digraph G, we can select a set *U(g)* of input vertices and a set *V(g)* of output vertices. Under the front top of the cluster is a vertex, which includes the outside edge of the cluster, under the output vertex refers to the vertex the edge goes to the outside of the cluster. Sets *U(g),V(g)* can intersect and can be empty. For simplicity and without loss of generality, assume that all the clusters of the graph $\mathcal{G}$, which are not part of the ribs and/or of which they don't have ribs, are unimodal.

Denote G the set of all clusters in the digraph G. Then the set of edges $Q^* = Q(\mathrm{G}) \setminus \bigcup_{g \in G} Q(g)$ consists of edges coming from the output vertices of one cluster to the input vertices of other clusters. Assume that we map each cluster g∈G to an acyclic digraph $\overline{g}$ that satisfies the inclusion $U(g) \cup V(g) \subseteq P(\overline{g})$.

Then, connecting the input and output vertices of acyclic digraphs with the edges of a set Q*, we obtain an acyclic digraph containing all the input and output vertices of clusters g∈G. In this case, an acyclic subgraph $\overline{g}$ in which there are no feedbacks, is built from each cluster, but all the links between the clusters remain. Replacing all clusters g∈G with acyclic subgraphs $\overline{g}$ can significantly simplify the appearance of the protein network presented in the form of a digraph G and in its formulation resembles the method of the main components in mathematical statistics.

Construction of acyclic digraph with input and output nodes of single cluster

Therefore, our task is to allocate in each cluster g∈G an acyclic subgraph $\overline{g}$ containing all vertices from sets $U(g), V(g)$: $U(g) \cup V(g) \subseteq P(\overline{g})$. The algorithm for transforming a cluster *g* into an acyclic digraph $\overline{g}$ consists of the following steps.

First, we construct an acyclic subgraph $\hat{g}$ of an digraph *g*, containing a non-empty subset $U(\hat{g}) \subseteq U(g)$ of the set of input vertices and the set *V(g)* of all output vertices. Then an acyclic subgraph $\tilde{g}$, is constructed consisting of paths starting at any vertex of the set and ending at any vertex of the set $U(\tilde{g}) = U(g) \setminus U(\hat{g})$ and ending at any vertex of the set $P(\hat{g})$. It is obvious that the union $\overline{g} = \hat{g} \cup \tilde{g}$ also forms an acyclic graph, and the inclusion $V(g) \subseteq P(\hat{g})$ is true.

Construction of acyclic graph $\hat{g}$

Construct an acyclic subgraph $\hat{g}*$ of digraph *g*, containing a non-empty subset $U(\hat{g}) \subseteq U(g)$ of the set of input vertices and the set *V(g)* of all output vertices. For this purpose, in the cluster g, we implement a recursive wave front algorithm from the set *U(g)* of all input vertices until the front passes through all output vertices *V(g)*:

$$A_0 = U(g),\ A_{k+1} = \left\{ j \notin \bigcup_{t=0}^{k} A_t : \exists\, i \in A_k, (i,j) \in Q(g) \right\},\ k = 0,\ldots,T-1, \quad (1)$$

$$T = \min\left\{ k : V(g) \subseteq \bigcup_{t=0}^{k} A_t \right\},\ P(\hat{g}*) = \bigcup_{t=0}^{T} A_t,\ V(g) \subseteq P(\hat{g}*).$$

From each output vertex $v \in V(g)$ that necessarily belongs to one of the constructed sets, for example $v \in A_{r(v)}$,the following recursive procedure along the graph $\hat{g}*$ reaches to the set of input vertices:

$$B_0^v = \{v\},\ B_{k+1}^v = \left\{ i \in A_{r(v)-k-1} : \exists\, j \in B_k^v,\ (i,j) \in Q(\hat{g}*) \right\},\ k = 0,\ldots,r(v)-1. \quad (2)$$

As a result, a subgraph $\hat{g}*(v)$ is constructed in the graph $\hat{g}*$. Then

***Corresponding author:** Tsitsiashvili Gurami, Institute for Applied Mathematics FEB RAS, Vladivostok, Russia, Tel: 49228734875; E-mail: guram@iam.dvo.ru

Received February 15, 2018; **Accepted** February 23, 2018; **Published** March 01, 2018

Citation: Gurami T, Victor B, Alexandr L (2018) Replacement of Directed Graph by Acyclic Directed Graph and Its Application in Biostatistics. J Biom Biostat 9: 390. doi: 10.4172/2155-6180.1000390

the subgraph $\hat{g} = \bigcup_{v \in V(g)} \hat{g}^*(v)$ of digraph $\hat{g}^*$ satisfies the following ratios $\bigcup_{v \in V(g)} B^v_{r(v)} = U(\hat{g}), \quad V(g) \subset P(\hat{g})$.

Construction of acyclic graph $\tilde{g}$

The construction of a graph $\tilde{g}$ is similar to the construction of a graph $\hat{g}$ with some changes. First, the wave front algorithm from the set $P(\hat{g})$ is implemented until the front passes $U(\tilde{g})$, through all the vertices of the set changing the orientation of the edges (i,j) in the ratio (1) to the opposite (j,i). Next, a recursive procedure (2) is implemented from each vertex $u \in U(\tilde{g})$, replacing the orientation of the edges (j,i) by (j,i).

Example

Represent graphically protein network, using for research of plants stability to drought and extreme temperatures, in accordance with previous study [6].

In Figure 1, clusters (containing more than single node) are outlining with blue and yellow curves. The clusters are constructing by algorithm described [7]. In Figure 2, we represent digraph in which

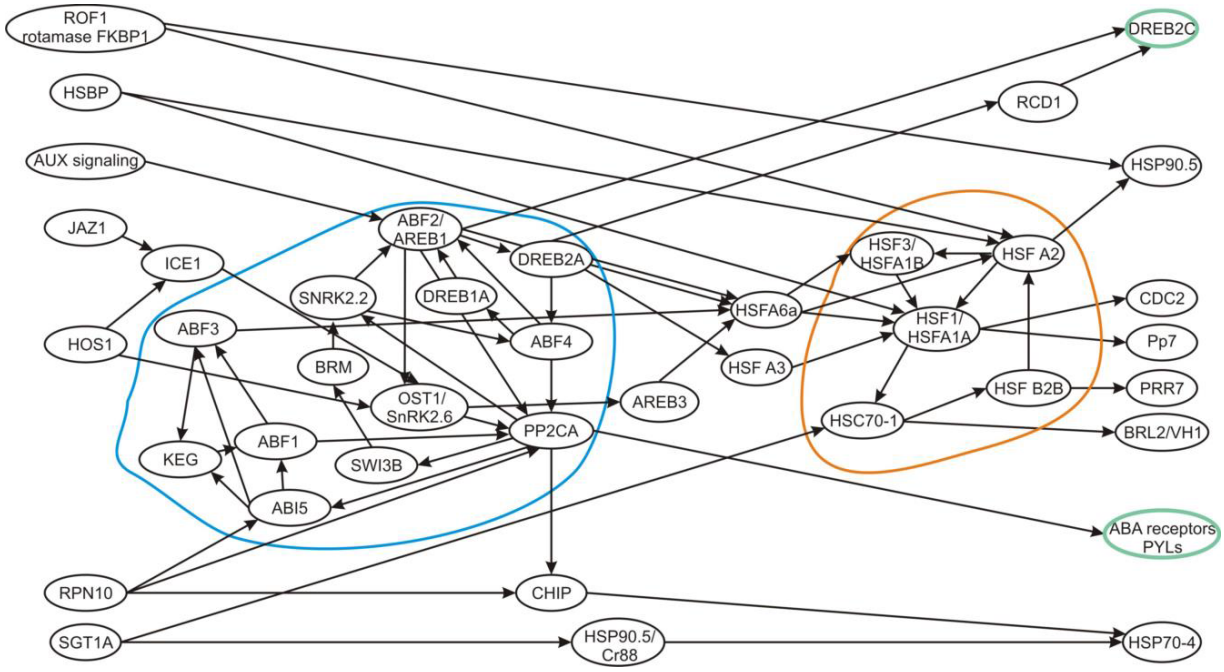

Figure 1: Factorization of protein network using for research of plants stability to drought and extreme temperatures.

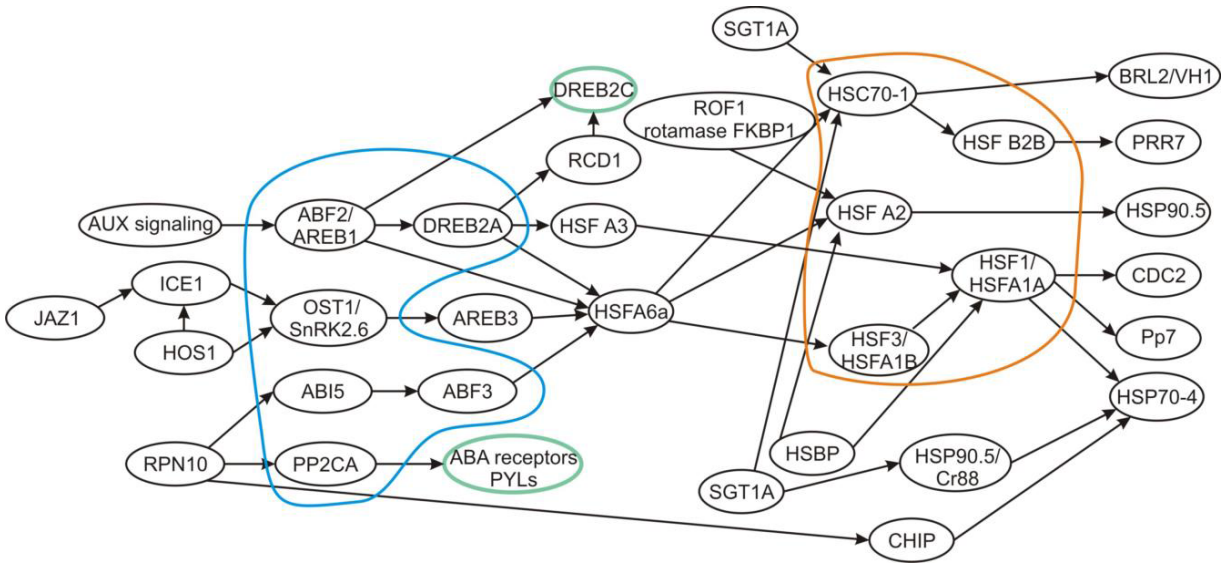

Figure 2: Replacement of clusters by shortest ways between input and output nodes in clusters.

each cluster is replacing by acyclic digraph. Next simplification is in an extraction from Figure 2 sub graph connected with nodes DREBC2 and ABA receptors PYLs, supplying thermo tolerance of the protein network only.

The transitions Figure 1 → Figure 2 → Figure 3 are basing on the suggested algorithms and their mathematical properties. Comparison of Figure 1 with Figure 3 shows possibilities of the network simplification for solution of considered biophysical problem.

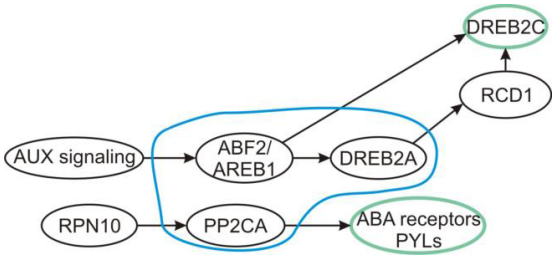

Figure 3: Extraction of ways from input nodes to nodes DREB2C and ABA receptors PYLs supplying thermo tolerance of the protein network.

Acknowledgement

Partially supported by Russian Fund of Basic Researches (project 17-07-00177).

References

1. Tsitsiashvili G (2017) Calculation of Flows Intensities Transformations in Acyclic Directed Networks. J Biom Biostat 8: 6.
2. Verslues PE (2016) ABA and cytokinins: challenge and opportunity for plant stress research. Plant Mol Biol 91: 629-640.
3. Vishwakarma K, Upadhyay N, Kumar N, Yadav G, Singh J, et al. (2017) Abscisic Acid Signaling and Abiotic Stress Tolerance in Plants: A Review on Current Knowledge and Future Prospects. Front Plant Sci 20: 8-161.
4. Bulgakov VP, Avramenko TV, Tsitsiashvili GS (2017) Critical analysis of protein signalling networks involved in the regulation of plant secondary metabolism: focus on anthocyanins. Crit Rev Biotechnol 37: 685- 700.
5. Tsitsiashvili GS, Bulgakov VP, Losev AS (2017) Factorization of Directed Graph Describing Protein Network, Using for Research of Plants Stability to Drought and Extreme Temperatures and its Possible Development. BBOAJ 3: 1.
6. Pircon N, Fortemaison CH, Jacobs S, Dremier J, Dumont E, et al. (2000) The visual display of regulatory information and networks. Trends in Cell Biology 10: 404-408.
7. Tsitsiashvili G (2013) Sequential algorithms of graph nodes factorization. Reliability: Theory & Applications 8: 30-33.