Additional files Information

Additional file 1: *The two networks representation*
1. Network files: Cytoscape and plain text format
Additional file 2: *Complex networks parameters*
2. A. Closeness centrality
2. B. Clustering coefficients
2. C. Betweenness centrality
2. D. Neighborhood connectivity (in and out)
2. E. Neighborhood connectivity (only out)
2. F. Neighborhood connectivity (only in)
2. G. In-degree distribution
2. H. Out-degree distribution
2. I. Shortest paths distribution
2. J. Stress centrality
2.K. Mutual Information Distribution for the edges of the 91 TF with increased connectivity in the tumor network
Additional file 3: *Full list of nodes that increase their activity*
3. A. List of TFs that most increase their activity in tumors
3. B. List of targets that most increase their activity in tumors
Additional file 4: *Genes with altered activity and mutations in COSMIC database*
4. A. TFs with increased activity and mutations COSMIC database
4. B. Targets with increased activity and mutations COSMIC database
4. C. TFs with decreased activity and mutations COSMIC database
4. D. Targets with decreased activity and mutations COSMIC database
Additional file 5: *Clusters enrichment analysis*
5. A. Cluster analysis of normal network and enrichment of KEGG pathways
5. B. Cluster analysis of tumor network and enrichment of KEGG pathways

Additional file 1: *The two networks representation*

1. Network files

The two networks are provided in diverse format for visualization: binary Cystoscape objects and plain text files.

In the Cystoscape objects the transcription factors and target genes were represented with the red and blue nodes, respectively. Furthermore the gene type (transcription factor or target) was stored as node attribute. Mutual information and edge direction were stored as edge attributes. Edges line width was correlated with the mutual information value for easy viewing.

In the plain text files the both networks were also provided as a simple list of edges.

- List of edges file: First line contains the attributes names. Then, each line contains the source transcription factor interaction, the target interaction and the mutual information value for this interaction.

- Types of nodes file: First line contains the attributes names. Then, each line contains the gene name and the gene type, "TF" for transcription factors and "Target" for target genes.

Additional file 2: *Complex networks parameters*

Full set of figures could be found in Additional file 2 (PDF document).

Additional file 3: *Full list of nodes that increase their activity*

3. A. List of TFs that most increase their activity in tumors

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 3.A).

3. B. List of targets that most increase their activity in tumors

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 3.B).

Column description for tables 3A and 3B

Gene.Symbol --> Approved HGNC gene symbol.

Entrez.Gene.ID --> Gene identifier of Entrez Gene database.

Normal.Outdegree --> Node out-degree en network of normal colon cells.

Tumor.Outdegree --> Node out-degree en network of tumor colon cells.

Difference.Outdegree --> Difference between Tumor.Outdegree and Normal.Outdegree.

Additional file 4: *Genes with altered activity and mutations in COSMIC database*

4. A. TFs with increased activity and mutations COSMIC database

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 4.A).

4. B. Targets with increased activity and mutations COSMIC database

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 4.B).

4. C. TFs with decreased activity and mutations COSMIC database

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 4.C).

4. D. Targets with decreased activity and mutations COSMIC database

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 4.D).

Column description for tables 4A-4D

Gene.Symbol --> Approved HGNC gene symbol.

Mutated.samples --> Number of unique mutated samples in COSMIC database for this gene.

Samples.tested --> Total number of samples tested in COSMIC database for the specified gene.

Percent --> Frequency of mutated samples in COSMIC database for this gene.

Additional file 5: *Clusters enrichment analysis*

5. A. Cluster analysis of normal network and enrichment of KEGG pathways

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 5.A).

5. B. Cluster analysis of tumor network and enrichment of KEGG pathways

Pathways colored in green are significantly enriched in at least one of the tumor network clusters but they do not appear in any cluster of the normal network. Pathways colored in red are common in both networks.

Large supplementary table submitted as MS Excel spreadsheet document (Sheet 5.B).

Column description for tables 5A and 5B

cluster --> Cluster ID.

pwyid --> Sigora pathway ID.

pwyname --> Name of KEGG pathway.

p_value --> P--value

adj.p_value --> Adjusted p--value (Bonferroni).

count --> Count of present GPS of the tested pathway.

weightsum --> Sum of the weights of all present GPS of the tested pathway.

round_sum (success) --> Rounded sum of the weights of all present GPS of the tested pathway.

m (sample size) --> Rounded sum of the weights of all possible GPS.

n (universe) --> Rounded sum of the weights of all possible GPS of the repository -- m column.

k (success states) --> Rounded sum of the weights of all possible GPS of the tested pathway.