Taking brain tumor as an example, we propose a new method, Mutual information gene networks - Structural key genes (MIGN - SKG), to identify key genes for dynamic biological system.

**Step 1:** Download the original data from <https://www.ncbi.nlm.nih.gov/>. Save data (gene expressions of samples for the control group and the experimental groups) as Excel files (see “Mutual\_Information\Workdatabases”)

**Step 2:** Run “Mutual\_lnformation\MI\_Calculator.m” in Matlab, then you will get the mutual information of genes in the corresponding dataset. If the dataset has N genes, then you will get N x N matrix as the result, where the element of the matrix in the Ath row and the Bth column is the mutual information between gene A and gene B (See “Mutual\_Information\MI\_Results”).

**Remark 1:** In the “Mutual\_lnformation\MI\_Calculator.m”, we use the following code to import the Excel file in Step 1:

"A=xlsread('file name(include the path)',' the sheet name','a specific rectangular region (range) in Excel to save the p-values of all genes ');"

When you run this file, please make the change corresponding to the Excel file you used.

**Remark 2:** Other m-files in the folder "Mutual\_lnformation " are the files (i.e., subfunctions) needed to successfully run “MI\_Calculator.m”. Specifically,

* “Discrete\_realdata.m” is to discretize the p- values in gene expression profiles.
* “Entroy\_realdata.m” is to calculate the Shannon entropy H (A) for gene A with the discretized p- value. For the original gene expression profiles, you could use “.Entroy.m” is to calculate its Shannon entropy.
* “Interentroy\_realdata.m” (“Interentroy.m”) is to calculate the Mutual Information I (A,B) using H(A) and H(B) obtained by “Entroy\_realdata.m” (“Entroy.m”).

**Step 3:** To select a uniform mutual information threshold to highlight the core structure of all MIGNs, we let the mutual information threshold change from 0.1 to 0.9 with step size 0.02 by using the code “threshold=0.1:0.02:0.9”. For a given MIGN (obtained in Step 2), there is a subnetwork corresponding to any fixed MI threshold. You could calculate network statistics of any mutual information gene subnetwork by running the following file:

“Networkstatistics\graph\_statistics.m”

The result will be a vector consisting of network average degree K, the ratio of non-isolated points R, average path length L, the average clustering coefficient C and the modularity Q. See “Networkstatistics\ MIGN-KLCQR”.

**Remark 3:** If you do not want to fix the step size, then you could import any MI threshold values saved in an Excel file by using the code:

threshold=xlsread ('file name(include the path)’,'the name of sheet','a specific rectangular region (range) in Excel to save MI thresholds');

**Remark 4:** In the above “graph\_statisitcs.m”, we import MIGN (MI matrix obtained in Step 2) by using the following code:

“load D:\matlab\work\Mutual\_Information\MI\_Results \I\_E1; ”

When you run "graph\_statistics.m", please modify the above code according to the location where you save the m file obtained in step 2.

**Remark 5:** Other m-files in the folder "Networkstatistics " are the files (i.e., subfunctions) needed to successfully run “graph\_statistics.m”. Specifically,

* “clusteringcoef.m”: compute C for the network
* “formatnet\_N11.m”: for any MI threshod and MI matrix, compute the non-isolated node and then establish mutual information gene subnet
* “newman\_N11”: compute Q for the network

**Step 4:** When the MI threshold θ exceeds a certain value, the modularity Q of those MIGNs begins to oscillate. It implies that the basic community structure of those gene networks changes. Therefore, according to Q values obtained in Step 3, you could select a uniform MI threshold θ and then obtain corresponding MI gene subnetworks , which are call the core MIGNs of the dynamic biology system.   
The analysis of the statistics obtained in Step 3 shows that K is the key network statistic that characterizes those MIGNs' structural differences. According to the change in the degree of nodes in the core MIGNs, you could identify four types of structural key genes (SKGs).