COPYRIGHT NOTICE

©2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

CITE THIS AS:

R. Gupta, S. M. Fayaz and S. Singh, "Identification of gene network motifs for cancer disease diagnosis," 2016 IEEE Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER), Mangalore, 2016, pp. 179-184.
doi: 10.1109/DISCOVER.2016.7806253

Identification of Gene Network Motifs for Cancer Disease Diagnosis

Rohit Gupta*, Fayaz S.M[†] and Sanjay Singh*

*Department of Information and Communication Technology

[†]Department of Biotechnology

Manipal Institute of Technology, Manipal University, Manipal-576104, India

Email:{rohit.gupta7@learner.manipal.edu,fayaz.shaik@manipal.edu,sanjay.singh@manipal.edu}

Abstract—All networks, including biological networks, computer networks, social networks and more can be represented as graphs, which include a number of small module such as sub-graph, also called as network motifs. Network motifs are sub-graph which recur themselves in a specific network or different networks. In biological networks, these network motifs plays very important role to identify diseases in human beings. In this paper we have developed a module to identify common network motifs types from cancer pathways and Signal Transduction Networks (STNs). It also identifies the topological behaviors of cancer networks and STNs. In this study, we have implemented five motif algorithms such as Auto-Regulation Loop (ARL), Feed Backward Loop (FBL), Feed Forward Loop (FFL), Single-Input Motif (SIM) and Bi-fan.

These algorithms gives correct results in terms of network motifs for human cancer and STNs. Finding network motifs by using online tool is limited to three nodes, but our proposed work provides facility to find network motifs up to any number of nodes. We applied five motif algorithms to human cancer networks and Signal Transduction Networks (STNs) which are collected from KEGG database as a result we got "Frequent Occurrences of Network Motifs (FONMs)". These FONMs acts as a references for an oncologist in order to find type of cancer in human being.

Keywords—Network Motifs, Cancer Network, STNs, Topological Relation

I. INTRODUCTION

In biological networks, basic interaction patterns that recur throughout the biological network, much more often than in random networks. The same small set of network motifs appears to serve as the building block of transcription networks from bacteria to mammals, neuronal networks and other biological and non biological networks.

Graph theory approach plays an important role for finding diseases in networks, based on their topological structures, such as protein-protein interaction network (PPIN) [1] and metabolic network (MN) [2][3]. Examples based on structures are: Auto-Regulation Loop (ARL), Feedback Loop (FBL), Feed-Forward Loop (FFL, coherent or incoherent), bi-fan and Single-Input Motif (SIM) [2]-[4]. These five network motifs play a important role in the form of molecular adjustment when some changes occurs like temperature, chemical concentration [5].

Previous studies says that these motifs are commonly found in organisms, such as the FFL is found in E.coli[6], in other bacteria [7], in yeast [8][9] and higher organisms

transcriptional regulatory network [10][11]; FBL and FFL motifs present in different types of network, such as neural networks, and PPIN [12][13]. It is claimed that network motifs do not have specific biological function, there is no significant behavior for network motifs [14], while other works shows different results [15][16].

Cancer is caused by genetic and epigenetic regulation. Sometimes because of alterations signaling pathways that control cell growth and division, cell death etc such changes leads to cancer of the STNs [17]. The origin of cancer is due to malfunctioning of genetic components of the STNs such as, MAPK, PI3K-Akt, Ras, jak-Stat etc. If once is affected, the signal will amplified and leads to cancer.

Rest of the paper is organized as follows. Section II briefly describes about genes, microRNA etc. Section III gives data collection, miRNAs regulation network motifs, proposed model, network motifs identification, motifs finding algorithms and topological relation of the network. Section IV gives results of cancer network motifs and STNs, miRNA, gene, TF motifs results and cancer network and STNs topological results and section V concludes this paper.

II. THEORETICAL BACKGROUND

This section gives brief background about genes, microRNAs, and transcription factors. A gene is the basic physical and functional unit of heredity. Genes comes from DNA (deoxyribonucleic acid), act as instruction to make proteins. In recent year, many research came up based on miRNA regulated network motifs. Micro-ribonucleic acid (miRNAs) are small endogenous RNAs of 21–25 nucleotides (nts) in length. They play an important regulatory role in animals and plants by targeting specific mRNAs for degradation or translation repression [18]. Moreover study reported miRNA regulation present in cell STNs [19]. Also some report tell miRNA expression is associated with tumor progression and metastasis. MiRNAs can be leads to cancer by targeting oncogenes (OCG) or tumor suppressor genes (TSG) [20][21]. It is also known that miRNA regulated network motifs for cancer network, which can be obtained from Kyoto Encyclopedia of Genes and Genomes (KEGG)[22].

Transcription factors (TFs) plays very important role to form gene regulatory network (GRN). In a recent work, Circuits DB [23] was publish, which gives FFL motif information from TFs, miRNAs and genes.

```
graph TD; Database[(Database)] --> KEGG[(KEGG)]; Database --> TAR[(TAR)]; Database --> ChipBase[(ChipBase)]; Database --> TARbase[(TARbase)]; KEGG --> NetworkMotifs[Network Motifs]; NetworkMotifs --> ARL[ARL]; NetworkMotifs --> FBL[FBL]; NetworkMotifs --> FFL[FFL]; NetworkMotifs --> SIM[SIM]; NetworkMotifs --> bifan[bifan]; TAR --> miRNARegulation[miRNA Regulation]; ChipBase --> miRNARegulation; TARbase --> miRNARegulation; miRNARegulation --> Gene[Gene]; miRNARegulation --> TF[TF]; miRNARegulation --> miRNA[miRNA];
```

result. In the graph theory approach, adjacency matrix having '1' that form the network. Based on getting network all five algorithms will run and get better result.

E. Motifs Finding Algorithms

1) *ARL*: This motifs type involves a self regulated gene. Non-zero diagonal elements in the adjacency matrix represent this type of motif, which is obtained using algorithm 1.

Algorithm 1: Finding occurrences of ARL present in the network A

Input : The network A with n nodes
Output: All ARL networks

```

1 for i= 1 to row do
2   for j= 1 to column do
3     if in the network A, the node of i and node of j
       are equal then then
4       Again check they having one in same
         position,increment counter until all network
         traverse ;
5     end
6   end
7 end
8 return counter

```

2) *FBL*: This motif type involves two genes regulate each other. For any location (i, j) in the adjacency matrix, if the term (i, j) is '1' and that of (j, i) is also '1', then genes i and j form a FBL, algorithms for finding FBL is given in algorithm 2.

Algorithm 2: Finding occurrences of FBL present in the network A

Input : The network A with n nodes
Output: All FBL networks

```

1 for i= 1 to row do
2   for j= 1 to column do
3     if in the network A, the position of (i,j) is one
       and (j,i) is also one, increment counter until all
       network traverse then
4     end
5   end
6 end
7 return counter

```

3) *FFL*: This motif type involves three genes regulating each other. Depending on the activation or suppression order, this motif type can be further divided into the so called cFFL, and iFFL.

For any triple set (i, j, k) , if the terms of (i, j) , (j, i) , (i, k) , (k, i) , (j, k) , and (k, j) are all of '1', then genes i , j , and k form a FFL, and the technique to identify this motif is given in algorithm 3.

4) *SIM*: SIM motif denotes a topology where a master gene regulates multiple downstream genes. Select any row in the adjacency matrix and count how many '1' appear in the row. Since there are at most n '1's in a row and n rows to search from algorithm 4.

Algorithm 3: Finding occurrences of FFL present in the network A

Input : The network A with n nodes
Output: All FFL networks

```

1 flag ← zero
2 for i= 1 to row do
3   for j= 1 to column do
4     if in the network A, (i,j) is one and (i,j), (j,i)are
       same then
5       print (i,j); for k=j+1 to row do
6         if in the network A, (k,j)is equal to one
           and (k,i)is equal to one and (k,i)(j,k)are
           same and(k,i)(i,k)are same,increase
           counter by one then
7           flag ← one
8           break;
9         end
10      end
11    end
12    if flag is equal to one then
13      end
14      break;
15    end
16  end
17 counter/=2;
18 print value of counter

```

Algorithm 4: Finding occurrences of SIM present in the network A

Input : The network A with n nodes
Output: All SIM networks

```

1 for i= 1 to row do
2   for j= 1 to column do
3     if in the network A, i and j is one, increment
       counter then
4     end
5   end
6   Again check counter is greater than one, increment
   count until all network traverse
7 end
8 return count

```

5) *Bi-fan*: Bi-fan motif denotes a topology where two genes regulate the same other two genes. To identify this network motif select any two rows in the adjacency matrix which have the value of '1' appear at the same column more than one time. Check whether these two rows are connected, if not, then determine which two columns have the value of '1' in both rows, algorithm 5 identifies Bi-fan type of network motifs.

F. Topological Relation of the Network

Complex networks may have topological relation, in our study there are 14 cancer networks and 14 STNs, and they form complex network structure. To find topologies between them we have used igraph package from R programming. Surprisingly result shows very good relation between the networks. We are using three parameters to analyze networks. These

Algorithm 5: Finding occurrences of Bifan present in the network A

Input : The network A with n nodes
Output: All Bifan networks

```

1 for i= 1 to row do
2   for j= i+1 to row do
3     for k=i to column do
4       for l=j to column do
5         if in the network A (i,k) is greater than
           one and (j,l) also greater than one then
6           if Again check in the network A (i,k)
             is equal to one or (l,j) is equal to
             one, and increment count until
             traverse all network A then
7             end
8           end
9         else
10          if In the network A if there are all
              one means (i,k),(i,l),(j,k) and (j,l)
              increment counter until reverse all
              network A then
11            end
12          end
13        end
14      end
15    end
16  end
17 return count

```

parameters are Degree Centrality (DC), Closeness Centrality (CC), and Betweenness Centrality (BC). Degree centrality is number of connection in the favorable nodes, it is defined as:

$$DC(i) = \frac{1}{N-1} \sum_j A_{ij} \quad (1)$$

where A_{ij} is the values in the adjacency matrix A and N is the total number of nodes in the network. $DC(i)$ is average over all i .

The Closeness Centrality is the total shortest possible line between two lines, from the center not too far, and it is defined as:

$$CC(A) = \left[\frac{1}{N-1} \sum_{j=1}^N d(A, j) \right]^{-1} \quad (2)$$

where A is the sum of the nodes from all other nodes, N is the total number of nodes in the network, d is the distance and j is the all connected node to A .

Betweenness is a centrality measure of a vertex inside a diagram. Vertices that happen on numerous shortest ways between different vertices have higher betweenness than those that don't.

For a graph $G=(V,E)$ with n vertices, the betweenness $BC(i)$ for vertex is computed as follows:

- 1) For each pair of vertices (j,k) , compute all shortest paths between them.
- 2) For every pair of vertices (j,k) , decide the division of shortest ways that go through the vertex.

- 3) Sum this fraction over all pairs of vertices (j,k) .

$$BC(i) = \sum_{j \neq k} d_{jk}(i) / d_{jk} \quad (3)$$

where $d_{jk}(i)$ is the number of shortest paths connecting jk passing through i and d_{jk} is the total number of shortest path.

IV. RESULTS AND DISCUSSION

A. Results of Cancer Network Motifs and STNs

TABLE I. CANCER MOTIFS RESULTS

Network	ARL	FBL	FFL	SIM	Bifan
AML	0	0	0	11	22
Basal Cell Carcinoma	0	0	0	1	0
Bladder Cancer	0	0	0	1	0
Chronic myeloid leukemia	0	0	0	4	1
Colorectal Cancer	0	0	0	4	3
Endometrial Cancer	0	0	0	4	1
Glioma	0	0	0	11	66
Melanoma	0	0	0	3	0
NSCLC	0	0	0	11	8
Pancreatic Cancer	0	1	1	10	9
Pathways in Cancer	0	0	0	33	189
PC	0	0	0	6	8
RCC	0	0	0	5	0
Small Cell Lung Cancer	0	0	0	5	0

TABLE II. SIGNAL TRANSDUCTION NETWORK (STNs)

Network	ARL	FBL	FFL	SIM	Bifan
Calcium Signalling	0	0	0	9	0
ErbB	0	0	0	21	64
FoxO	0	0	0	6	6
Hedgehog	0	0	0	3	7
HIF1	0	0	0	6	1
Hippo	0	0	0	10	5
Jak-Stat	0	0	0	2	13
Mapk	0	2	2	36	46
mTOR	0	0	0	7	8
NFKB	0	0	0	10	2
Notch	0	0	0	2	6
Phosphatidylinositol	0	74	0	37	36
PI3k-Akt	0	0	0	13	2
Rap1	0	0	0	11	6

A total 14 cancer network and 14 STNs were used. Table I gives the cancer network motifs result and Table II gives STNs motifs result. From Table I and II, we observe that the FOMNs of motifs are SIM and bi-fan, but rarely ARL, FBL, and FFL motifs are present. We note that overall FOMNs in cancer and STNs based on the database selection it may vary. In our process and selection of database gives result like SIM and bifans is FOMNs. Based on the adjacency matrix, it allow us to identify exact number of motifs, there is no need of p values[33].

B. miRNA, Gene, TF Motifs Results

TABLE III. THREE COMPONENT NETWORK

From TarBase, ChipBase and TAG Database					
	ARL	FBL	FFL	SIM	Bifan
Network Motifs	0	0	0	62	2768

Figure 2 shows that the controlling gene in the regulation of overall biological function are AKT1, ABL1, AKT3, ABL2, ANP34A, AKAP13, APC and more. These gene are important

in forming relation between the biological process. In Table III there are three component, interaction we find all type of motifs from adjacency matrix concept and here also found that bifan and SIM motifs are more often in cancer network and STNs. So we can say that this type of pattern occurring in the cancer network is more often.

C. Cancer Networks and STNs Topological Results

Five different types of network motifs given in Table I and II. SIM and bifan motifs are the dominant among all type of motifs. We have analyzed network structure using DC , CC and BC as a parameter to understand the internal structure of both the network.

TABLE IV. CANCER MOTIFS TOPOLOGICAL RESULTS

Network	DC	CC	BC
AML	0.125	0.038	0.301
Basal Cell Carcinoma	0.199	0.038	0.121
Bladder Cancer	0.061	0.010	0.037
Chronic myeloid leukemia	0.060	0.016	0.111
Colorectal Cancer	0.074	0.014	0.092
Endometrial Cancer	0.095	0.017	0.090
Glioma	0.041	0.005	0.049
Melanoma	0.051	0.019	0.075
NSCLC	0.081	0.021	0.205
Pancreatic Cancer	0.082	0.009	0.076
Pathways in Cancer	0.036	0.002	0.095
PC	0.191	0.024	0.301
RCC	0.129	0.008	0.054
Small Cell Lung Cancer	0.129	0.008	0.054
Median, α	1.294	0.229	1.661

TABLE V. SIGNAL TRANSDUCTION NETWORK(STNs) TOPOLOGICAL RESULTS

Network	DC	CC	BC
Calcium Signalling	0.250	0.027	0.504
ErbB	0.146	0.030	0.204
FoxO	0.282	0.005	0.116
Hedgehog	0.309	0.014	0.106
HIF1	0.337	0.019	0.398
Hippo	0.066	0.002	0.019
Jak-Stat	0.289	0.048	0.402
Mapk	0.092	0.020	0.360
mTOR	0.193	0.036	0.337
NFKB	0.125	0.007	0.177
Notch	0.307	0.088	0.479
Phosphatidylinositol	0.110	0.026	0.223
PI3k-Akt	0.178	0.016	0.458
Rap1	0.182	0.005	0.182
Median, β	2.866	0.343	3.965

TABLE VI. RATIO OF TOPOLOGICAL RESULTS

	Topological Results		
	DC	BC	CC
Ratio, γ	2.214	1.497	2.387

Let α be the median of all cancer network and β be the median of all STNs. γ is the ratio of α and β . From the Table IV we found the cancer network median α for all topological parameter and from Table V we found STNs median β of all topological parameter. Table VI gives the ratio of cancer networks and STNs which tells DC, CC and BC is higher for STN as compared to cancer network. This is due to the fact that STNs have more dense gene-gene interaction, and it appears to be very important component of the STNs network. Apart from that the chances of getting the cancer is more.

V. CONCLUSION

In this paper, we have used adjacency matrix concept to finding the network motifs and proposed motif identification module, which identifies the network motif of any size. Our results says that bi-fan and SIM motifs are two most frequently found motifs in cancer networks and STNs. And we found that the topological behavior of the network also plays important role in cancer diagnosis.

We have developed a motif identification module which may help oncologist to do the quick objective analysis of cancer data for a patient and come to conclusion about the cancer type a patient is suffering with, and take required interventions. We are able to show this by taking the different database of the same network, we found bi-fans and SIM motifs are often occurring pattern in the cancer network and STNs.

REFERENCES

- [1] A. Wagner, "The Yeast Protein Interaction Network Evolves Rapidly and Contains Few Redundant Duplicate Genes," *Molecular Biology and Evolution*, vol. 18, no. 7, pp. 1283–1292, Jul. 2001. [Online]. Available: <http://mbe.oxfordjournals.org/content/18/7/1283.abstract>
- [2] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási, "The large-scale organization of metabolic networks," *Nature*, vol. 407, pp. 651–654, 2000. [Online]. Available: <http://www.nature.com>
- [3] H. Ma and A. P. Zeng, "Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms," *Bioinformatics*, vol. 19, pp. 270–277, 2003.
- [4] A. Ma'ayan, S. L. Jenkins, R. L. Webb, S. I. Berger, S. P. Purushothaman, N. S. Abul-Husn, J. M. Posner, T. Flores, and R. Iyengar, "SNAVI: Desktop application for analysis and visualization of large-scale signaling networks," *BMC systems biology*, vol. 3, no. 1, 2009. [Online]. Available: <http://dx.doi.org/10.1186/1752-0509-3-10>
- [5] T. M. Yi, Y. Huang, M. I. Simon, and J. Doyle, "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proc Natl Acad Sci USA*, vol. 97, no. 9, pp. 4649–53, Apr. 2000. [Online]. Available: <http://www.pnas.org/content/97/9/4649.long>
- [6] A. Lipshtat, S. P. Purushothaman, R. Iyengar, and A. Ma'ayan, "Functions of bifans in context of multiple regulatory motifs in signaling networks," *Biophysical Journal*, vol. 94, no. 7, pp. 2566 – 2579, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0006349508705113>
- [7] P. Eichenberger, M. Fujita, S. T. Jensen, E. M. Conlon, D. Z. Rudner, S. T. Wang, C. Ferguson, K. Haga, T. Sato, J. S. Liu, and R. Losick, "The Program of Gene Transcription for a Single Differentiating Cell Type during Sporulation in *Bacillus subtilis*," *PLoS Biol*, vol. 2, no. 10, pp. e328+, Sep. 2004. [Online]. Available: <http://dx.doi.org/10.1371/journal.pbio.0020328>
- [8] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon, "Network Motifs: Simple building blocks of complex networks," *Science*, vol. 298, pp. 824–827, 2002. [Online]. Available: <http://www.sciencemag.org/cgi/content/abstract/sci.298/5594/824>
- [9] T. I. Lee, N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J.-B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford, and R. A. Young, "Transcriptional Regulatory Networks in *Saccharomyces cerevisiae*," *Science*, vol. 298, no. 5594, pp. 799–804, Oct. 2002. [Online]. Available: <http://dx.doi.org/10.1126/science.1075090>
- [10] D. T. Odom, N. Zizlsperger, D. B. Gordon, G. W. Bell, N. J. Rinaldi, H. L. Murray, T. L. Volkert, J. Schreiber, P. A. Rolfe, D. K. Gifford, E. Fraenkel, G. I. Bell, and R. A. Young, "Control of Pancreas and Liver Gene Expression by HNF Transcription Factors," *Science*, vol. 303, no. 5662, pp. 1378–1381, Feb. 2004. [Online]. Available: <http://dx.doi.org/10.1126/science.1089769>

- [11] N. Iranfar, D. Fuller, and W. F. Loomis, "Transcriptional regulation of post-aggregation genes in dictyostelium by a feed-forward loop involving {GBF} and lagc," *Developmental Biology*, vol. 290, no. 2, pp. 460 – 469, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0012160605008602>
- [12] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. S. Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, "Superfamilies of evolved and designed networks." *Science*, vol. 303, no. 5663, pp. 1538–1542, Mar. 2004. [Online]. Available: <http://www.sciencemag.org/cgi/content/abstract/303/5663/1538>
- [13] A. Ma'ayan, S. L. Jenkins, S. Neves, A. Hasseldine, E. Grace, B. Dubin-Thaler, N. J. Eungdamrong, G. Weng, P. T. Ram, J. J. Rice, A. Kershenbaum, G. A. Stolovitzky, R. D. Blitzer, and R. Iyengar, "Formation of regulatory patterns during signal propagation in a Mammalian cellular network." *Science (New York, N.Y.)*, vol. 309, no. 5737, pp. 1078–1083, Aug. 2005. [Online]. Available: <http://dx.doi.org/10.1126/science.1108876>
- [14] P. Ingram, M. Stumpf, and J. Stark, "Network motifs: structure does not determine function," *BMC Genomics*, vol. 7, no. 1, pp. 108+, 2006. [Online]. Available: <http://dx.doi.org/10.1186/1471-2164-7-108>
- [15] S. Mangan and U. Alon, "Structure and function of the feed-forward loop network motif," *Proceedings of the National Academy of Sciences*, vol. 100, no. 21, pp. 11980–11985, Oct. 2003. [Online]. Available: <http://dx.doi.org/10.1073/pnas.2133841100>
- [16] G. Swiers, R. Patient, and M. Loose, "Genetic regulatory networks programming hematopoietic stem cells and erythroid lineage specification," *Developmental Biology*, vol. 294, no. 2, pp. 525 – 540, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0012160606001679>
- [17] R. Sever and J. S. Brugge, "Signal transduction in cancer," *Cold Spring Harbor perspectives in medicine*, vol. 5, no. 4, p. a006098, 2015.
- [18] D. P. Bartel, "MicroRNAs: Genomics, biogenesis, mechanism, and function," *Cell*, vol. 116, no. 2, pp. 281 – 297, 2004. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0092867404000455>
- [19] E. Wang, *RNA Technologies in Cardiovascular Medicine and Research*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, ch. MicroRNA Systems Biology, pp. 69–86. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-78709-9_5
- [20] L. He, J. M. Thomson, M. T. Hemann, E. Hernando-Monge, D. Mu, S. Goodson, S. Powers, C. Cordon-Cardo, S. W. Lowe, G. J. Hannon *et al.*, "A microRNA polycistron as a potential human oncogene," *nature*, vol. 435, no. 7043, pp. 828–833, 2005.
- [21] R. Garzon, M. Fabbri, A. Cimmino, G. A. Calin, and C. M. Croce, "MicroRNA expression and function in cancer," *Trends in Molecular Medicine*, vol. 12, no. 12, pp. 580 – 587, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1471491406002425>
- [22] A. Nakaya, T. Katayama, M. Itoh, K. Hiranuka, S. Kawashima, Y. Moriya, S. Okuda, M. Tanaka, T. Tokimatsu, Y. Yamanishi, A. C. Yoshizawa, M. Kanehisa, and S. Goto, "KEGG OC: a large-scale automatic construction of taxonomy-based ortholog clusters," *Nucleic Acids Research*, vol. 41, no. D1, pp. D353–D357, Jan. 2013. [Online]. Available: <http://dx.doi.org/10.1093/nar/gks1239>
- [23] O. Friard, A. Re, D. Taverna, M. De Bortoli, and D. Corá, "CircuitsDB: a database of mixed microRNA/transcription factor feed-forward regulatory circuits in human and mouse." *BMC bioinformatics*, vol. 11, no. 1, pp. 435+, 2010. [Online]. Available: <http://dx.doi.org/10.1186/1471-2105-11-435>
- [24] A. Nakaya, T. Katayama, M. Itoh, K. Hiranuka, S. Kawashima, Y. Moriya, S. Okuda, M. Tanaka, T. Tokimatsu, and e. a. Y. Yamanishi, "Kegg oc: a large-scale automatic construction of taxonomy-based ortholog clusters," *Nucleic acids research*, vol. 41, pp. D353–357, 2013.
- [25] JS.Chen, WS.Hung, HH.Chan, SJ.Tsai, and H. Sun, "In silico identification of oncogenic potential of fyn-related kinase in hepatocellular carcinoma," *Bioinformatics*, vol. 29, pp. 420–427, 2013.
- [26] JH.Yang, JH.Li, S.Jiang, H.Zhou, and LH.Qul, "Chipbase: a database for decoding the transcriptional regulation of long non-coding rna and microrna genes from chip-seq data," *Nucleic acids research*, vol. 41, pp. D177–187, 2013.
- [27] SD.Hsu, YT.Tseng, S.Shrestha, Y. A.Khaleel, CH.Chou, CF.Chou, HY.Huang, C. Lin, SY.Ho, and et al, "mirtarbase update 2014: an information resource for experimentally validated mirna-target interactions," *Nucleic acids research*, vol. 24, pp. D78–84, 2013.
- [28] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: a software environment for integrated models of biomolecular interaction networks." *Genome Research*, vol. 13, pp. 2498–2504, Nov. 2003.
- [29] F. Schreiber and H. Schwöbbermeyer, "MAVisto: a tool for the exploration of network motifs." *Bioinformatics*, Jul. 2005. [Online]. Available: <http://view.ncbi.nlm.nih.gov/pubmed/16020473>
- [30] S. Wernicke and F. Rasche, "Fanmod: A tool for fast network motif detection," *Bioinformatics*, vol. 22, no. 9, pp. 1152–1153, May 2006. [Online]. Available: <http://dx.doi.org/10.1093/bioinformatics/btl038>
- [31] A. Ferro, R. Giugno, G. Pigola, A. Pulvirenti, D. Skripin, G. D. Bader, and D. Shasha, "Netmatch: a cytoscape plugin for searching biological networks," *Bioinformatics*, vol. 23, pp. 910–912, 2007.
- [32] A. Ma'ayan, "Insights into the Organization of Biochemical Regulatory Networks Using Graph Theory Analyses," *Journal of Biological Chemistry*, vol. 284, no. 9, pp. 5451–5455, Feb. 2009. [Online]. Available: <http://dx.doi.org/10.1074/jbc.r800056200>
- [33] S. Wernicke and F. Rasche, "Fanmod: a tool for fast network motif detection," *Bioinformatics*, vol. 22, no. 9, pp. 1152–1153, 2006. [Online]. Available: <http://bioinformatics.oxfordjournals.org/content/22/9/1152.abstract>