World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:15, No:5, 2021

Identifying Network Subgraph-Associated Essential Genes in Molecular Networks

Efendi Zaenudin, Chien-Hung Huang, Ka-Lok Ng

***Abstract*—**Essential genes play an important role in the survival of an organism. It has been shown that cancer-associated essential genes are genes necessary for cancer cell proliferation, where these genes are potential therapeutic targets. Also, it was demonstrated that mutations of the cancer-associated essential genes give rise to the resistance of immunotherapy for patients with tumors. In the present study, we focus on studying the biological effects of the essential genes from a network perspective. We hypothesize that one can analyze a biological molecular network by decomposing it into both three-node and four-node digraphs (subgraphs). These network subgraphs encode the regulatory interaction information among the network's genetic elements. In this study, the frequency of occurrence of the subgraph-associated essential genes in a molecular network was quantified by using the statistical parameter, odds ratio. Biological effects of subgraph-associated essential genes are discussed. In summary, the subgraph approach provides a systematic method for analyzing molecular networks and it can capture useful biological information for biomedical research.

***Keywords*—**Biological molecular networks, essential genes, graph theory, network subgraphs

I. Introduction

ESSENTIAL genes are genes that are necessary and sufficient to maintain the survival of an organism. Experimentally, the biologists systematically knocked out each of the functional genes in an organism to explore whether the organism can survive or reproduce. This group of essential genes make up the so-called minimal genome [1].

It has been shown that cancer-associated essential genes are those necessary for cancer cell proliferation [2], [3], where these genes are potential therapeutic targets [4]. It was shown that the presence of essential genes near a deletion cancer essential gene may decrease the frequency of homozygous deletions [5]. Also, it was demonstrated that mutations of the cancer-associated essential genes give rise to resistance of immunotherapy for tumor patients [6]. Furthermore, disease genes which play the role of essential genes are more likely than non-essential genes resistant to therapeutic treatments [7].

We noted that most of the studies focus on collecting essential gene information [8] and prediction [4], in the present study, we focus on studying the essential genes on a network perspective.

Based on our previous study [9] we assume that both 3-node and 4-node network directed subgraphs (digraphs) are the fundamental building blocks of a molecular network. We do not consider randomized versions of the studied network. We have determined that the number of both 3-node and 4-node network subgraph patterns are the same as the number of both 3-node and 4-node network motif patterns. As a matter of fact, there are a total of 13 and 199 possible patterns can be defined for both 3-node and 4-node subgraphs (motifs), respectively. In other words, subgraphs are treated as the core network components. This is similar to the work of Mowshowitz [10], who proposed that a finite graph (V vertices and E edges) can be decomposed into equivalence classes (h classes).

In a recent work [11], we extended the previous work [12] by developing a subgraph identification tool named *PatternFinder* to identify both 3-node and 4-node subgraphs in cancer networks, signal transduction networks, and cellular processes.

In the 'Method' section, we give a description of the three input datasets and the methods used in this paper. In the 'Results' section, results for the association of network subgraphs and essential genes are reported. We conclude in the 'Conclusions' section.

II. Methods

A. Input Datasets

In this study, we selected molecular networks from the KEGG database [13]. KEGG provides a comprehensive collection of molecular network information that were prepared in the KGML format (August 2017). Not every network recorded by KEGG was retrieved in this work. We removed networks composed of several disconnected subnetworks with repetitive regulatory structures ("Two-component system" and "MicroRNAs in cancer"), small networks with size less than 10 ("Chemical carcinogenesis" and "Viral carcinogenesis"). In addition, we collected the networks labeled with the name "signaling pathway," and called them "signal transduction networks (STNs)".

Cancer networks are collected from two families in the KEGG classification: "Cancer: overview," and "Cancer: specific types". We note that STNs range across different families in the KEGG classification, including "Signal transduction," "Immune system," "Endocrine system," and

Efendi Zaenudin is with the Asia University, Taichung 41354, Taiwan and Research Center for Informatics, Indonesian Institute of Sciences, Indonesia (e-mail: ezaenudin@mail.informatika.lipi.go.id). EZ is supported by the grant of Ministry of Science and Technology of Taiwan (MOST) MOST 108-2221-E- 468-020 and 109-2221-E-468-013.

Chien-Hung Huang is with the National Formosa University, Yunlin 632, Taiwan (e-mail: chhuang@nfu.edu.tw). CHH is supported by the grants: NFU EN2019_10141052529786 and MOST 109-2221-E-150-036.

Ka-Lok Ng is associated with the Asia University, Taichung 41354, Taiwan and Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 404, Taiwan (corresponding author, phone: +886-4-2339-4541; e-mail: ppiddi@gamil.com). KLN is supported by the grants: MOST 108-2221-E- 468-020 and 109-2221-E-468-013, and grants from Asia University, 105-asia-11, 106-asia-06, 106-asia-09, 107-asia-02 and 107-asia-09.

"Nervous system". Cellular processes cover three different families from the KEGG database, i.e. "Cell growth and death," "Cellular community - eukaryotes," and "Cell motility".

A total of 71 networks were collected, i.e. 17 cancer networks, 45 STNs, and 9 cellular processes. We downloaded KEGG pathway KGML files and utilized the KEGGScape [14] and KEGGparser [15] Cytoscape plug-in tools to get the node and edge information for each network.

OEGGv2 [2] is a database which collected experimentally tested essential, fitness genes (intermediate essentiality statuses) and non-essential genes. We selected essential genes (*Homo sapiens*) from OEGGv2, and because the essential genes have an alias name, we therefore included an alias (gene ID) using the *Bioconductor* package "org.Hs.egALIAS2EG" (DOI: 10.18129/B9.bioc.org.Hs.eg.db). Hence, a total of 227 genes were collected from the OGEEv2 database (http://ogee.medgenius.info).

B. Network Connectivity - Adjacency Matrix

By integrating the nodes and edges information from KEGGScape and KEGGparser, using a MATLAB code, we constructed an adjacency matrix, $\boldsymbol{A}$, to represent each one of the 71 networks. The typical sizes of the networks are around 100, whereas, real world molecular networks compose of thousands of genes, which are much larger than the networks we analyzed; however, the regulatory and feedback interactions (digraphs) information among thousands of genes are not available in KEGG at the present time.

C. Network Subgraphs Detection Tool

In total, there are 13 3-node subgraphs and 199 4-node subgraphs [16], [17]. Each subgraph can be represented by an adjacency matrix; hence it can be converted into a decimal. We have developed an algorithm named *PatternFinder* to enumerate all possible 3-node subgraphs and 4-node subgraphs embedded in the 71 networks. Details about *PatternFinder* are given in [9] Supplementary File 1 – Supplementary Table S4.

D. Network Subgraph-Associated Essential Genes

We propose to examine network subgraph-associated essential genes for cancer networks, STN and cellular processes. Given a molecular network, we consider the 2x2 contingency table (Table I), which depicts the statistics of essential genes and non-essential genes embedded in 'subgraph module' and 'non-subgraph module'.

TABLE I
THE 2X2 CONTINGENCY TABLE FOR ESSENTIAL GENES AND NON-ESSENTIAL GENES EMBEDDED IN 'SUBGRAPH MODULE' AND 'NON-SUBGRAPH MODULE'

	Subgraph module	Non-subgraph module	Total
Essential genes	*a*	*b*	*a* + *b*
Non-essential genes	*c*	*d*	*c* + *d*
Total	*a* + *c*	*b* + *d*	*a* + *b*+ *c* + *d*

Subgraph module of a network is given by the union of the 3-node subgraphs and 4-node subgraphs. We used a statistical parameter, odds ratio (*OR*), to estimate the level of propensity of essential genes found in a subgraph module. *OR* measures the relative odds of finding essential genes embedded in network subgraph modules relative to non-subgraph modules. The *OR* is defined by:

$$\mathrm{OR} = \frac{\frac{\mathrm{p(essential\_gene \mid subgraph\_module)}}{1-\mathrm{p(essential\_gene \mid subgraph\_module)}}}{\frac{\mathrm{p(essential\_gene \mid non\_subgraph\_module)}}{1-\mathrm{p(essential\_gene \mid non\_subgraph\_module)}}} = \frac{a \times d}{b \times c} \quad (1)$$

If *OR* > 1, it means that network subgraphs are enriched with essential genes.

III. RESULTS

The results of the association of network subgraphs and essential genes for cancer networks, STNs, and cellular processes are given in Table II in the 'Appendix' section. The *OR* are listed in the last column. An *OR* greater than one indicates that essential genes are enriched in the subgraph module (consisting of 3-node and 4-node subgraph genes, but the subgraphs are not necessarily interconnected). Among the 17 cancer networks, none of the network has an *OR* greater than 1. For the 45 STNs, three have an *OR* larger than 1, i.e. Hippo signaling pathway (*OR* = 1.636), NF-kappa B signaling pathway (*OR* = 1.343) and TNF signaling pathway (*OR* = 2.615). Among the 9 cellular processes, one has an *OR* greater than 1, i.e. Signaling pathways regulating pluripotency of stem cells (*OR* = 1.514). Odds ratios indicate that these four networks are enriched with subgraph-associated essential genes.

The Hippo signaling pathway composes of five essential genes; i.e. 'BIRC2', 'BIRC5', 'FGF1', 'PPP2CA', and 'SMAD1'. The Hippo signaling pathway is a newly found pathway that is involved in cancer development [18] and controlling organ size [19] and angiogenesis [20].

The NF-kappa B signaling pathway consists of three essential genes, i.e. 'TNF', 'XIAP', and 'BIRC2'. This pathway plays an important role in cancer [21], [22], inflammatory diseases [22], metabolic diseases [23] and rheumatoid arthritis [24]. The TNF signaling pathway consists of three essential genes, i.e. 'TNF', 'LTA', and 'BIRC2'. TNF-alpha is a cytokine involved in systemic inflammation, plays an important role in homeostasis, immunity and participates in hematopoietic stem cell survival and myeloid regeneration [25]. Furthermore, the TNF family cytokines activate both canonical and noncanonical NF-kappa B pathways [26] through TNFR1 and CD40, respectively [27].

Signaling pathways regulate pluripotency of stem cells; i.e. 'SMAD1' and 'FGF2'. Using the CRISPR-Cas9 screening technique, a large-scale study has been conducted to chart the essential genes for human pluripotent stem cells [28].

We noted that among the 71 networks we studied, the cell cycle network consists of the largest number of subgraph-associated essential genes (the OR is less than one). The cell cycle process consists of eight essential genes, i.e. 'ESPL1', 'PCNA', 'CDC45', 'PLK1', 'WEE2', 'CDC6', 'PRKDC' and 'CHEK1'. Yu et al. reported that mutations in essential genes disrupted cell cycle progression in the yeast [29]. This result is expected because essential genes are highly involved with the

survivability of a cell.

IV. Conclusions

In conclusion, this study provides a systematic approach to dissecting the underlying structure of biological molecular networks. The use of network subgraph approach serves as a powerful technique to dissect the underlying topology in terms of 3-node and 4-node subgraphs. The next step is to test our hypothesis by analyzing 5-node subgraphs [30]. We expect that our efforts will further elucidate the biological nature of molecular networks.

Appendix

TABLE II
The Results of the Association of Essential Genes and the 3-Node and 4-Node Subgraph Module and Non-Subgraph Module

Cancer Networks		Genes embedded in subgraph	Genes not embedded in subgraph	Odds ratio/total
Acute myeloid leukemia [hsa05221]	Essential gene	1	0+1 (pseudo-count)	0.024
	Non-essential gene	42	0+1 (pseudo-count)	
	Total	43	2	45
Basal cell carcinoma [hsa05217]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.095
	Non-essential gene	21	2	
	Total	22	3	25
Breast cancer [hsa05224]	Essential gene	2	0+1 (pseudo-count)	0.490
	Non-essential gene	49	12	
	Total	51	13	64
Choline metabolism in cancer [hsa05231]	Essential gene	1	0+1 (pseudo-count)	0.385
	Non-essential gene	26	10	
	Total	27	11	37
Chronic myeloid leukemia [hsa05220]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.270
	Non-essential gene	37	10	
	Total	38	11	49
Colorectal cancer [hsa05210]	Essential gene	1	1	0.278
	Non-essential gene	36	10	
	Total	37	11	48
Endometrial cancer [hsa05213]	Essential gene	1	0+1 (pseudo-count)	0.286
	Non-essential gene	28	8	
	Total	29	9	38
Gastric cancer [hsa05226]	Essential gene	2	0+1 (pseudo-count)	0.840
	Non-essential gene	50	21	
	Total	52	22	74
Glioma [hsa05214]	Essential gene	1	0+1 (pseudo-count)	0.167
	Non-essential gene	30	5	
	Total	31	6	37
Hepatocellular carcinoma [hsa05225]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.547
	Non-essential gene	53	29	
	Total	54	30	84
Melanoma [hsa05218]	Essential gene	1	0+1 (pseudo-count)	0.125
	Non-essential gene	24	3	
	Total	25	4	29
Non-small cell lung cancer [hsa05223]	Essential gene	1	0+1 (pseudo-count)	0.2
	Non-essential gene	35	7	
	Total	36	8	44
Pancreatic cancer [hsa05212]	Essential gene	1	1	0.067
	Non-essential gene	45	3	
	Total	46	4	50
Pathways in cancer [hsa05200]	Essential gene	6	0+1 (pseudo-count)	0.431
	Non-essential gene	153	11	
	Total	159	12	171
Prostate cancer [hsa05215]	Essential gene	1	0+1 (pseudo-count)	0.486
	Non-essential gene	35	17	
	Total	36	18	54
Renal cell carcinoma [hsa05211]	Essential gene	0+1 (pseudo-count)	1	0.344
	Non-essential gene	29	10	
	Total	29	11	41
Small cell lung cancer [hsa05222]	Essential gene	2	0+1 (pseudo-count)	0.706

Cancer Networks		Genes embedded in subgraph	Genes not embedded in subgraph	Odds ratio/total
	Non-essential gene	34	12	
	Total	36	13	49
STN		Genes embedded in subgraph	Genes not embedded in subgraph	Odds ratio/total
Adipocytokine signaling pathway [hsa04920]	Essential gene	1	0+1 (pseudo-count)	0.094
	Non-essential gene	32	3	
	Total	33	4	37
AMPK signaling pathway [hsa04152]	Essential gene	1	1	0.333
	Non-essential gene	45	15	
	Total	46	16	61
Apelin signaling pathway [hsa04371]	Essential gene	1	0+1 (pseudo-count)	0.174
	Non-essential gene	46	8	
	Total	47	9	56
B cell receptor signaling pathway [hsa04662]	Essential gene	0+1 (pseudo-count)	1	0.324
	Non-essential gene	34	11	
	Total	35	12	47
Calcium signaling pathway [hsa04020]	Essential gene	1	0+1 (pseudo-count)	0.870
	Non-essential gene	23	20	
	Total	24	21	45
cAMP signaling pathway [hsa04024]	Essential gene	0+1 (pseudo-count)	1	0.090
	Non-essential gene	67	6	
	Total	68	7	75
cGMP-PKG signaling pathway [hsa04022]	Essential gene	1	0+1 (pseudo-count)	0.208
	Non-essential gene	48	10	
	Total	49	11	60
Chemokine signaling pathway [hsa04062]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.063
	Non-essential gene	48	3	
	Total	49	4	53
C-type lectin receptor signaling pathway [hsa04625]	Essential gene	2	0+1 (pseudo-count)	0.086
	Non-essential gene	70	3	
	Total	72	4	76
ErbB signaling pathway [hsa04012]	Essential gene	2	0+1 (pseudo-count)	0.040
	Non-essential gene	50	1	
	Total	52	2	54
Estrogen signaling pathway [hsa04915]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.133
	Non-essential gene	30	4	
	Total	31	5	36
Fc epsilon RI signaling pathway [hsa04664]	Essential gene	0+1 (pseudo-count)	2	0.133
	Non-essential gene	30	8	
	Total	31	10	41
FoxO signaling pathway [hsa04068]	Essential gene	2	0+1 (pseudo-count)	0.212
	Non-essential gene	66	7	
	Total	68	8	76
Glucagon signaling pathway [hsa04922]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.452
	Non-essential gene	31	14	
	Total	32	15	47
GnRH signaling pathway [hsa04912]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.111
	Non-essential gene	36	4	
	Total	37	5	42
Hedgehog signaling pathway [hsa04340]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.043
	Non-essential gene	23	0+1 (pseudo-count)	
	Total	24	2	25
HIF-1 signaling pathway [hsa04066]	Essential gene	3	1	0.551
	Non-essential gene	49	9	
	Total	52	10	62
Hippo signaling pathway [hsa04390]	Essential gene	5	0+1 (pseudo-count)	1.636
	Non-essential gene	55	18	
	Total	60	19	79
Insulin signaling pathway [hsa04910]	Essential gene	1	1	0.130
	Non-essential gene	54	7	

Cancer Networks		Genes embedded in subgraph	Genes not embedded in subgraph	Odds ratio/total
Jak-STAT signaling pathway [hsa04630]	Total	55	8	63
	Essential gene	1	0+1 (pseudo-count)	0.031
	Non-essential gene	32	1	
MAPK signaling pathway [hsa04010]	Total	33	2	35
	Essential gene	2	0+1 (pseudo-count)	0.072
	Non-essential gene	111	4	
mTOR signaling pathway [hsa04150]	Total	113	5	118
	Essential gene	1	0+1 (pseudo-count)	0.200
	Non-essential gene	55	11	
Neurotrophin signaling pathway [hsa04722]	Total	56	12	68
	Essential gene	1	0+1 (pseudo-count)	0.043
	Non-essential gene	70	3	
NF-kappa B signaling pathway [hsa04064]	Total	71	4	74
	Essential gene	3	1	1.343
	Non-essential gene	67	30	
NOD-like receptor signaling pathway [hsa04621]	Total	70	31	101
	Essential gene	3	0+1 (pseudo-count)	0.581
	Non-essential gene	93	18	
Notch signaling pathway [hsa04330]	Total	96	19	115
	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.588
	Non-essential gene	17	10	
Oxytocin signaling pathway [hsa04921]	Total	18	11	29
	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.130
	Non-essential gene	46	6	
p53 signaling pathway [hsa04115]	Total	47	7	54
	Essential gene	3	0+1 (pseudo-count)	0.429
	Non-essential gene	49	7	
Phosphatidylinositol signaling system [hsa04070]	Total	52	8	60
	Essential gene	1	0+1 (pseudo-count)	0.115
	Non-essential gene	26	3	
Phospholipase D signaling pathway [hsa04072]	Total	27	4	31
	Essential gene	1	0+1 (pseudo-count)	0.195
	Non-essential gene	41	8	
PI3K-Akt signaling pathway [hsa04151]	Total	42	9	51
	Essential gene	2	1	0.597
	Non-essential gene	67	20	
PPAR signaling pathway [hsa03320]	Total	69	21	90
	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.109
	Non-essential gene	46	5	
Prolactin signaling pathway [hsa04917]	Total	47	6	53
	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.026
	Non-essential gene	38	1	
Rap1 signaling pathway [hsa04015]	Total	39	2	41
	Essential gene	1	0+1 (pseudo-count)	0.231
	Non-essential gene	65	15	
Ras signaling pathway [hsa04014]	Total	66	16	82
	Essential gene	1	0+1 (pseudo-count)	0.047
	Non-essential gene	64	3	
Relaxin signaling pathway [hsa04926]	Total	65	4	69
	Essential gene	1	0+1 (pseudo-count)	0.125
	Non-essential gene	48	6	
RIG-I-like receptor signaling pathway [hsa04622]	Total	49	7	56
	Essential gene	0+1 (pseudo-count)	1	0.957
	Non-essential gene	33	22	
Sphingolipid signaling pathway [hsa04071]	Total	34	23	56
	Essential gene	2	0+1 (pseudo-count)	0.512
	Non-essential gene	43	11	
T cell receptor signaling pathway [hsa04660]	Total	45	12	57
	Essential gene	0+1 (pseudo-count)	1	0.294

Cancer Networks		Genes embedded in subgraph	Genes not embedded in subgraph	Odds ratio/total
	Non-essential gene	51	15	
	Total	52	16	68
TGF-beta signaling pathway [hsa04350]	Essential gene	3	2	0.485
	Non-essential gene	34	11	
	Total	37	13	50
Thyroid hormone signaling pathway [hsa04919]	Essential gene	1	0+1 (pseudo-count)	0.333
	Non-essential gene	48	16	
	Total	49	17	66
TNF signaling pathway [hsa04668]	Essential gene	3	1	2.615
	Non-essential gene	39	34	
	Total	42	35	77
Toll-like receptor signaling pathway [hsa04620]	Essential gene	2	0+1 (pseudo-count)	0.219
	Non-essential gene	64	7	
	Total	66	8	74
VEGF signaling pathway [hsa04370]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.036
	Non-essential gene	28	0+1 (pseudo-count)	
	Total	29	2	31
Wnt signaling pathway [hsa04310]	Essential gene	2	0+1 (pseudo-count)	0.094
	Non-essential gene	64	3	
	Total	66	4	70
Cellular Processes		Genes embedded in subgraph	Genes not embedded in subgraph	Odds ratio/total
Adherens junction [hsa04520]	Essential gene	0+1 (pseudo-count)	0+1 (pseudo-count)	0.182
	Non-essential gene	44	8	
	Total	45	9	53
Apoptosis [hsa04210]	Essential gene	17	0+1 (pseudo-count)	0.355
	Non-essential gene	3	9	
	Total	79	10	89
Cell cycle [hsa04110]	Essential gene	8	10	0.432
	Non-essential gene	50	27	
	Total	58	37	95
Cellular senescence [hsa04218]	Essential gene	2	0+1 (pseudo-count)	0.576
	Non-essential gene	66	19	
	Total	68	20	88
Focal adhesion [hsa04510]	Essential gene	3	0+1 (pseudo-count)	0.053
	Non-essential gene	57	1	
	Total	60	2	62
Gap junction [hsa04540]	Essential gene	1	0+1 (pseudo-count)	0.030
	Non-essential gene	33	0+1 (pseudo-count)	
	Total	34	2	36
Necroptosis [hsa04217]	Essential gene	2	0+1 (pseudo-count)	0.246
	Non-essential gene	57	7	
	Total	59	8	67
Regulation of actin cytoskeleton [hsa04810]	Essential gene	0+1 (pseudo-count)	1	0.196
	Non-essential gene	56	11	
	Total	57	12	69
Signaling pathways regulating pluripotency of stem cells [hsa04550]	Essential gene	2	0+1 (pseudo-count)	1.514
	Non-essential gene	37	28	
	Total	39	29	68

A pseudo-count '1' was added if the number of embedded essential gene is zero, it is because the odds ratio (OR) is not well-defined. ORs and the total number of genes are listed in the last column. An OR greater than one implies that essential genes are enriched in the 3-node and 4-node subgraph module.

REFERENCES

[1] Juhas, M., L. Eberl, and J.I. Glass, *Essence of life: essential genes of minimal genomes.* Trends Cell Biol, 2011. 21(10): p. 562-8.

[2] Chen, W.-H., et al., *OGEE v2: an update of the online gene essentiality database with special focus on differentially essential genes in human cancer cell lines.* Nucleic Acids Research, 2016. 45(D1): p. D940-D944.

[3] Zhan, T. and M. Boutros, *Towards a compendium of essential genes - From model organisms to synthetic lethality in cancer cells.* Critical reviews in biochemistry and molecular biology, 2016. 51(2): p. 74-85.

[4] Gilvary, C., et al., *A machine learning approach predicts essential genes and pharmacological targets in cancer*. 2019, bioRxiv.

[5] Pertesi, M., et al., *Essential genes shape cancer genomes through linear limitation of homozygous deletions.* Communications Biology, 2019. 2(1): p. 262.

[6] Patel, S.J., et al., *Identification of essential genes for cancer immunotherapy.* Nature, 2017. 548(7669): p. 537-542.

[7] Dickerson, J.E., et al., *Defining the role of essential genes in human disease.* PloS one, 2011. 6(11): p. e27368-e27368.

[8] Zhang, R., H.Y. Ou, and C.T. Zhang, *DEG: a database of essential genes.* Nucleic Acids Research, 2004. 32(suppl_1): p. D271-D272.

[9] Huang, C.-H., et al., *Computational analysis of molecular networks using*

spectral graph theory, complexity measures and information theory. bioRxiv, 2019: p. 536318.
[10] Mowshowitz, A., *Entropy and the complexity of graphs: II. The information content of digraphs and infinite graphs.* The bulletin of mathematical biophysics, 1968. 30(2): p. 225-240.
[11] Lee, C.H. Huang, and K.L. Ng, *In silico study of significant network motifs in the cancer networks.* Master's thesis, National Formosa University, Taiwan., 2016.
[12] Hsieh, W.T., et al., *Transcription factor and microRNA-regulated network motifs for cancer and signal transduction networks.* BMC Syst Biol, 2015. 9 Suppl 1: p. S5.
[13] Nakaya, A., et al., *KEGG OC: a large-scale automatic construction of taxonomy-based ortholog clusters.* Nucleic Acids Res, 2013. 41(Database issue): p. D353-7.
[14] Nishida, K., et al., *KEGGscape: a Cytoscape app for pathway data integration.* F1000Res, 2014. 3: p. 144.
[15] Arakelyan, A. and L. Nersisyan, *KEGGParser: parsing and editing KEGG pathway maps in Matlab.* Bioinformatics, 2013. 29(4): p. 518-9.
[16] Alon, U., *An Introduction to Systems Biology: design principles of biological circuits.* 2006: Chapman and Hall/CRC.
[17] Shen-Orr, S.S., et al., *Network motifs in the transcriptional regulation network of Escherichia coli.* Nature Genetics, 2002. 31(1): p. 64-68.
[18] Chan, S.W., et al., *The Hippo pathway in biological control and cancer development.* J Cell Physiol, 2011. 226(4): p. 928-39.
[19] Pan, D., *Hippo signaling in organ size control.* Genes Dev, 2007. 21(8): p. 886-97.
[20] Boopathy, G.T.K. and W. Hong, *Role of Hippo Pathway-YAP/TAZ Signaling in Angiogenesis.* Frontiers in Cell and Developmental Biology, 2019. 7(49).
[21] Karin, M., et al., *NF-kappaB in cancer: from innocent bystander to major culprit.* Nat Rev Cancer, 2002. 2(4): p. 301-10.
[22] Park, M.H. and J.T. Hong, *Roles of NF-κB in Cancer and Inflammatory Diseases and Their Therapeutic Approaches.* Cells, 2016. 5(2).
[23] Baker, R.G., M.S. Hayden, and S. Ghosh, *NF-κB, inflammation, and metabolic disease.* Cell Metab, 2011. 13(1): p. 11-22.
[24] Sabir, J.S.M., et al., *Dissecting the Role of NF-κb Protein Family and Its Regulators in Rheumatoid Arthritis Using Weighted Gene Co-Expression Network.* Frontiers in Genetics, 2019. 10(1163).
[25] Yamashita, M. and E. Passegué, *TNF-α Coordinates Hematopoietic Stem Cell Survival and Myeloid Regeneration.* Cell Stem Cell, 2019. 25(3): p. 357-372.e7.
[26] Sun, S.-C., *Non-canonical NF-κB signaling pathway.* Cell research, 2011. 21(1): p. 71-85.
[27] Hayden, M.S. and S. Ghosh, *Regulation of NF-κB by TNF family cytokines.* Seminars in immunology, 2014. 26(3): p. 253-266.
[28] Yilmaz, A., et al., *Defining essential genes for human pluripotent stem cells by CRISPR-Cas9 screening in haploid cells.* Nat Cell Biol, 2018. 20(5): p. 610-619.
[29] Yu, L., et al., *A survey of essential gene function in the yeast cell division cycle.* Molecular biology of the cell, 2006. 17(11): p. 4736-4747.
[30] Zaenudin, E., et al., *A Parallel Algorithm to Generate Connected Network Motifs* IAENG International Journal of Computer Science, 2019 46(4): p. pp518-523.