



# Integrated Network Analysis Reveals FOXM1 and MYBL2 as Key Regulators of Cell Proliferation in Non-small Cell Lung Cancer

Firoz Ahmed<sup>1,2\*</sup>

<sup>1</sup> Department of Biochemistry, University of Jeddah, Jeddah, Saudi Arabia, <sup>2</sup> University of Jeddah Center for Scientific and Medical Research, University of Jeddah, Jeddah, Saudi Arabia

**Background:** Loss of control on cell division is an important factor for the development of non-small cell lung cancer (NSCLC), however, its molecular mechanism and gene regulatory network are not clearly understood. This study utilized the systems bioinformatics approach to reveal the “driver-network” involve in tumorigenic processes in NSCLC.

**Methods:** A meta-analysis of gene expression data of NSCLC was integrated with protein-protein interaction (PPI) data to construct an *NSCLC network*. MCODE and iRegulone were used to identify the local clusters and its upstream transcription regulators involve in NSCLC. Pair-wise gene expression correlation was performed using GEPIA. The survival analysis was performed by the Kaplan-Meier plot.

**Results:** This study identified a local “driver-network” with highest MCODE score having 26 up-regulated genes involved in the process of cell proliferation in NSCLC. Interestingly, the “driver-network” is under the regulation of TFs FOXM1 and MYBL2 as well as miRNAs. Furthermore, the overexpression of member genes in “driver-network” and the TFs are associated with poor overall survival (OS) in NSCLC patients.

**Conclusion:** This study identified a local “driver-network” and its upstream regulators responsible for the cell proliferation in NSCLC, which could be promising biomarkers and therapeutic targets for NSCLC treatment.

**Keywords:** non-small cell lung cancer, gene expression, meta-analysis, systems bioinformatics, gene network

## OPEN ACCESS

### Edited by:

Umberto Malapelle,  
University of Naples Federico II, Italy

### Reviewed by:

Francesco Pepe,  
University of Naples Federico II, Italy  
Francesco Passiglia,  
Palermo University Hospital, Italy

### \*Correspondence:

Firoz Ahmed  
fahmed1@uj.edu.sa

### Specialty section:

This article was submitted to  
Thoracic Oncology,  
a section of the journal  
Frontiers in Oncology

**Received:** 26 June 2019

**Accepted:** 20 September 2019

**Published:** 15 October 2019

### Citation:

Ahmed F (2019) Integrated Network Analysis Reveals FOXM1 and MYBL2 as Key Regulators of Cell Proliferation in Non-small Cell Lung Cancer. *Front. Oncol.* 9:1011. doi: 10.3389/fonc.2019.01011

## INTRODUCTION

Lung cancer is one of the most commonly diagnosed cancer with high mortality around the world (1). The global prevalence of lung cancer and mortality rate is rising at an alarming pace with an estimated number of newly diagnosed lung cancer was 2.1 million while the number of deaths was 1.8 million in 2018 (<https://gco.iarc.fr>). Based upon histology, lung cancer is divided into two classes: (i) Non-small cell lung cancer (NSCLC) which represents approximately 85–90% of all lung cancer, and (ii) Small-cell lung cancer (SCLC) which represents approximately 10–15% of the lung cancer (1). NSCLC has three major sub-classes including (a) lung squamous cell carcinoma (LUSC), (b) lung adenocarcinoma (LUAD), and (c) large cell carcinoma. However, due to lack of clinical symptom and effective diagnostic screening, the NSCLC is generally diagnosed at an advanced stage. The 5-year overall

survival rate of metastatic NSCLC is 6% and has not been significantly improved in spite of having advancement in surgery, chemotherapy, and radiation (<https://www.cancer.org>).

Molecular profiling of NSCLC identified mutations in the tumor suppressor genes (TP53, RB1), oncogenes (EGFR, KRAS, AKT, MAPK) and translocations in oncogenes (ALK, RET, ROS1, NTRK1, NRG1), which alter the important signal-transduction pathways (2, 3). EGFR mutants have been reported more frequently in NSCLC in nonsmokers Asians patients and showed highly sensitive to therapy with EGFR tyrosine kinase inhibitors such as gefitinib and erlotinib (4, 5). Similarly, ALK rearranged gene fusion was also highly reported in NSCLC and has been proven more effective treatment with ALK-targeted inhibitors (crizotinib and alectinib) (1, 6). The genomic mutations not only alter the protein structure but also affect the expression level of genes involved in the cell division resulting in uncontrolled cell proliferation, cell survival, and NSCLC. Previous studies mainly focused on understanding the alteration in gene expression in NSCLC tumors (7), and identified overexpressed genes including CDC20 (8), CCNB1 (9), ASPM (10), and KIF4A (11), which contributes to the proliferation of tumor cells and also associated with poor prognosis. Furthermore, the role of transcription factors (TFs) including MYBL2 (12), FOXM1 (13–15), and E2F4 (16) in cell proliferation and cell survival in NSCLC has been reported. The miRNAs, a class of small non-coding RNAs which regulate gene expression at the post-transcriptional level through binding to 3'UTR of mRNA (17, 18), are also emerging as promising biomarkers for detecting NSCLC (19). However, the molecular mechanism and regulatory network of the mRNAs, TFs, miRNAs, and proteins underlying dysregulated cell division and cell proliferation in NSCLC are still largely remain unclear. Addressing these challenges are most pivotal for developing anticancer drugs and diagnostic and prognostic biomarkers for better management and personalized treatment of NSCLC.

The emergence of high-throughput genomics, transcriptomics, proteomics, and interactome data and their integrative analysis opens a new avenue for a deep understanding of etiology of cancer (20, 21). This work is focused on applying a systems bioinformatics approach to uncover interaction and regulatory mechanism of mRNAs, TFs, miRNAs, and proteins underlying cell proliferation and progression of NSCLC. Gene expression profiles have been integrated to identify the high confidence up- and down-regulated genes in the NSCLC compared to adjacent non-tumor tissues. Moreover, using the transcriptome-interactome data, NSCLC network was constructed and analyzed to understand the molecular mechanism underlying the development and proliferation of NSCLC. Our analysis revealed one important “driver-network” consists of 26 genes and its upstream regulators TFs FOXM1 and MYBL2 whose overexpression are associated with dysregulation of cell cycle and enhance cell proliferation in NSCLC. Furthermore, NSCLC associated miRNAs regulating the genes of “driver-network” were also identified. Combination of genes in the “driver-network” and upstream regulators could be potential biomarkers for diagnosis and prognosis; and therapeutic targets for better treatment of NSCLC.

## MATERIALS AND METHODS

### Gene Expression Data Collection

In February 2019, microarray gene expression data were searched in Gene Expression Omnibus database (GEO: [www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) using following criteria: (a) Lung cancer; (b) Human; and (c) Expression profiling by array; which gave 304 unique GEO series (GSEs). Then, a careful manually selected the GSEs data using following criteria: (d) Each GSE must have the profile of NSCLC along with adjacent non-tumor tissues as a control; (e) Each group (NSCLC/control) must have more than 20 samples; (f) All GSEs are from same microarray platform. Based upon the above criteria, three GSEs data [GSE27262 (22, 23), GSE18842 (7), and GSE19804 (24)] were selected and downloaded for further study (Table S1).

### Identification of Differentially Expressed Genes (DEGs)

This study analyzed the gene expression data having 131 samples from NSCLC and 130 samples from adjacent non-tumor tissues as control (Normal). Preprocessing of each microarray raw data including background correction, normalization and log<sub>2</sub> transformation were performed separately with RMA of Oligo package version 1.46 in Bioconductor/R version 3.5.2 (25). Each normalized expression data was integrated into a single file and batch effects were removed with ComBat of sva package version 3.30 in R (26). After that, differential expression analysis of genes between NSCLC compared to control was calculated using the linear modeling features of the limma package version 3.38 in Bioconductor/R (27). Affymetrix probe set ids were mapped to gene symbol using DAVID 6.8 (<https://david.ncifcrf.gov/>) (28). The gene is considered as differentially expressed (DEGs) if log<sub>2</sub> Fold Change |log<sub>2</sub>FC| is >2 and adjusted *P*-value is < 0.001. If multiple probe id mapped with the same gene, probe id with highest log<sub>2</sub>FC were selected. The expression data of the significant DEGs were selected and transformed into Z-score (row-wise of value), then a hierarchical clustering across rows were performed to create a heatmap using Morpheus tool (<https://software.broadinstitute.org/morpheus/>).

### Functional Annotation and Pathway Enrichment Analysis

In order to investigate the biological processes altered in NSCLC, we performed the functional annotation including Gene Ontology (GO) enrichment analysis for Biological Process, Molecular Function, Cellular Component, and Kyoto Encyclopedia of Genes and Genomes (KEGG: [www.kegg.jp](http://www.kegg.jp)) to the list of DEGs. All these functional annotations were performed with clusterProfiler v3.10.1 in Bioconductor/R using *pvalueCutoff* = 0.01, *pAdjustMethod* = “BH,” *qvalueCutoff* = 0.05, *minGSSize* = 5 (29).

### Construction and Analysis of the NSCLC Network

To construct the NSCLC network, DEGs were mapped to the STRING version 11 application (30). The setting parameters of STRING were: (a) meaning of network edges (confidence);

(b) active interaction sources (selected all); (c) minimum required interaction score (high confidence >0.900); (d) max number of interactors to show (1st shell, none/query protein only); (2nd shell, none). The PPI data was downloaded, and the network was visualized in Cytoscape software version (3.5.1), where each node represents a protein while the edge represents an interaction between proteins (31). We also integrated the information of differential gene expression level into the network, where the red node indicates up-regulated, while the green node indicates down-regulated expression in NSCLC compared to control. The topological structure of the NSCLC network was analyzed using Cytoscape plug-in “NetworkAnalyzer.”

Intending to identify highly connected local sub-networks in the NSCLC network, we applied the Cytoscape plug-in MCODE clustering algorithm (32). Furthermore, the biological relevance of these modules was analyzed with GO and KEGG pathway.

## Identifying the Upstream Regulator of Genes in Cluster

To identify the upstream transcription regulators of genes in MCODE clusters, Cytoscape plug-in iRegulone (V 1.3) was used at default parameters (33). Then, in house program was used to generate a matrix table where each row indicates TF while a column indicates a target gene. A Venn diagram was drawn using (<http://www.interactivenn.net/>).

## Expression Correlation Between Genes and TFs

To establish the relationship between genes in a cluster and its upstream regulators, pair-wise gene expression correlation was performed with GEPIA (<http://gepia.cancer-pku.cn/index.html>) (34). The web server integrated RNA-seq expression data from 9,736 tumors and 8,587 normal samples of the Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) projects. The analysis was done on default parameters: Pearson correlation coefficient; and selected LUAD and LUSC as TCGA tumor and TCGA normal dataset.

## Effect of Signature Genes on Survival in NSCLC by Kaplan-Meier Plot

The potential effect of expression of relevant genes on the overall survival (OS) was analyzed on the lung cancer patient. An online KM plotter software (<http://kmplot.com/analysis/>) was used to generate the Kaplan-Meier Plot on 1926 NSCLC cancer patients (LUAD and LUSC) (35). The tool run on the default parameters on hazard ratio (HR) with 95% confidence intervals and log-rank *P*-value which is considered as significant *P*-value < 0.05. The biased arrays (*n* = 2,435) were excluded for quality control.

## Extension of Cluster 1 With miRNA

In order to understand the regulatory role of miRNAs in NSCLC, the differentially expressed miRNAs (DEMs) in NSCLC compared to normal was downloaded from miRCancer database (<http://mircancer.ecu.edu/>) (36). After removing redundancy and cleaning of the data, 56 miRNAs were appeared as

up-regulated, while 168 miRNAs were appeared as down-regulated in NSCLC compared to control. The targets of these DEMs were identified using miRNet tool (<https://www.mirnet.ca/>). After that, only those miRNAs were selected for further study which are targeting any of the 31 genes (*Cluster 1* and its associated TFs). Finally, an extended sub-network of *Cluster 1* was generated by integrating *Cluster 1* with its upstream regulators of TFs and miRNAs in Cytoscape. To make sparse visualization of network, interaction within *Cluster 1* as well as between *Cluster 1* and TFs were removed.

## Mutational Signatures in NSCLC

The cBio Cancer Genomic Portal (<http://cbiportal.org>) is a freely available tool to explore cancer genomic data in diverse cancers. We selected the NSCLC from TCGA database and submitted the list of 31 genes from *Cluster 1* and its associated TFs in cBioPortal.

## RESULTS

### Verification of Each Group of Samples Using Principal Component Analysis

The Principal Component Analysis (PCA) was performed on normalized data of gene expression, which revealed a clear difference between NSCLC and normal samples in each GSEs study (Figures S1A–C). The cumulative contribution of PC1, PC2, and PC3 is 38.59, 33.64, and 36.47% for GSE27262, GSE18842, and GSE19804 datasets, respectively. In order to increase the statistical power to discover the DEGs, the expression data of three GSEs was integrated

**TABLE 1** | List of top 20 differentially expressed genes in NSCLC.

Gene	Up-regulated		Down-regulated		
	Log2FC	Adj. <i>p</i> -value	Gene	Log2FC	Adj. <i>p</i> -value
SPP1	4.70	4.49E-66	AGER	-5.06	2.18E-89
COL11A1	4.27	3.81E-51	CLDN18	-5.05	4.65E-54
COL10A1	4.10	1.62E-67	SFTPC	-4.53	9.44E-41
MMP12	3.95	3.41E-47	GPM6A	-4.50	6.61E-89
MMP1	3.86	1.32E-35	ADH1B	-4.34	5.76E-51
GREM1	3.71	3.70E-43	FABP4	-4.30	8.44E-65
HS6ST2	3.56	4.04E-54	TMEM100	-4.28	8.16E-57
GJB2	3.39	3.21E-46	CLIC5	-4.15	2.31E-69
CTHRC1	3.38	6.08E-62	CA4	-4.13	1.45E-85
TOP2A	3.36	4.37E-66	FAM107A	-4.12	2.82E-77
ANLN	3.25	2.25E-59	WIF1	-4.03	1.35E-40
COL1A1	3.13	9.35E-49	FCN3	-4.02	2.07E-62
PSAT1	3.05	4.96E-59	GKN2	-3.90	1.01E-56
TMPRSS4	3.00	8.65E-52	STXBP6	-3.88	9.95E-66
SPINK1	2.93	1.21E-20	CD36	-3.88	2.41E-64
CDCA7	2.90	2.04E-56	Mt1m	-3.87	2.33E-50
CST1	2.90	6.08E-42	AQP4	-3.76	2.70E-37
CXCL14	2.87	5.54E-31	SFTPA1	-3.70	2.40E-28
CEACAM5	2.83	2.54E-23	cpb2	-3.69	1.29E-48
RRM2	2.76	8.65E-52	TNNC1	-3.69	4.28E-84

and technical variability and noise were removed using batch effect correction. The PCA analysis clearly showed the distinction between cancer and normal samples indicating successful removal of the batch effects on the GSEs microarray data (Figures S1D,E).

## Identification of DEGs

To identify the genes contributing to the NSCLC, differential expression analysis was conducted on gene expression data. A total of 346 DEGs including 97 up-regulated and 249 down-regulated genes were identified with  $|\log_2FC| > 2$  and adjusted  $P$ -value is  $< 0.001$ . Among DEGs, top 10 genes showing up-regulated expression are SPP1, COL11A1, COL10A1, MMP12, MMP1, GREM1, HS6ST2, GJB2, CTHRC1, and TOP2A; while top 10 genes showing down-regulated expression are AGER, CLDN18, SFTPC, GPM6A, ADH1B, FABP4, TMEM100, CLIC5, CA4, and FAM107A (Table 1). Detail information and the complete list of DEGs is provided in Table S2A for up-regulated and Table S2B for down-regulated genes. A hierarchical cluster heatmap of DEGs across biological samples reveals distinct patterns of gene expressions in NSCLC compared to normal (Figure S2).

## Functional Annotation and Pathway Enrichment Analysis

To understand the biological function and key pathways altered in NSCLC, function annotation and pathway enrichment analysis was performed for the list of up- and down-regulated genes. Biological process (BP) and Molecular Function (MF) of Gene Ontology analysis revealed that the up-regulated genes are primarily associated with nuclear division, organelle fission, chromosome segregation, regulation of mitotic nuclear division, metaphase/anaphase transition of cell cycle, and mitotic spindle assembly checkpoint, and microtubule binding (Figure 1A). The cellular components (CC) of up-regulated genes were significantly associated with spindle, chromosome region, kinetochore, microtubule, and midbody, fibrillar collagen trimer, and spindle microtubule (Figure 1A). The KEGG pathway analysis showed the up-regulated genes were significantly enriched in only Cell cycle-G2/M transition (Figure 1A).

Biological process (BP) and Molecular Function (MF) of Gene Ontology analysis revealed that the down-regulated genes are primarily associated with circulatory system process, leukocytes migration, cell-substrate adhesion, regulation of angiogenesis, receptor-mediated endocytosis, cell chemotaxis, regulation of cell junction assembly, amide binding, peptide binding, and cytokine binding (Figure 1B). The cellular components (CC) of down-regulated genes were significantly associated with the extracellular matrix, membrane raft, cell-cell junction, and collagen-containing extracellular matrix (Figure 1B). The KEGG pathway analysis showed the down-regulated genes were significantly enriched in only BMP signaling (Figure 1B). The complete results of GO and KEGG analyses could be found in Table S3A for up-regulated and Table S3B for down-regulated genes.

## Construction and Analysis of NSCLC Network

Mapping of DEGs on STRING gave PPI network with 151 nodes and 640 edges, which were visualized in Cytoscape software, where each node represents a protein while an edge represents an interaction between proteins. The gene expression level of each protein was integrated into the PPI network, where the red node indicates up-regulated, while the blue node indicates down-regulated gene expression level in NSCLC compared to normal and termed as *NSCLC network* (Figure 2). Size of the node is based upon the degree of connectivity of the node. Edges in the network represent direct interactions between nodes. As shown in Figure 2, there are 61 and 86 nodes in the network showing up-regulation and down-regulation, respectively; while 4 nodes are not having gene expression level (identified by PPI interaction and not in the list of our DEGs). A highly interconnected sub-network of overexpressed genes could be seen in the NSCLC network. The structural topological of *NSCLC network* including Betweenness Centrality, Closeness Centrality, Clustering Coefficient, and Degree were analyzed and presented in Table S4. Furthermore, highly inter-connected 15 sub-network clusters were extracted from *NSCLC network* using Cytoscape plug-in MCODE (Figure S3; Table 2). Among them, top five clusters with the highest MCODE score were considered for further study. Topologically relevant information of a gene is given as follows:

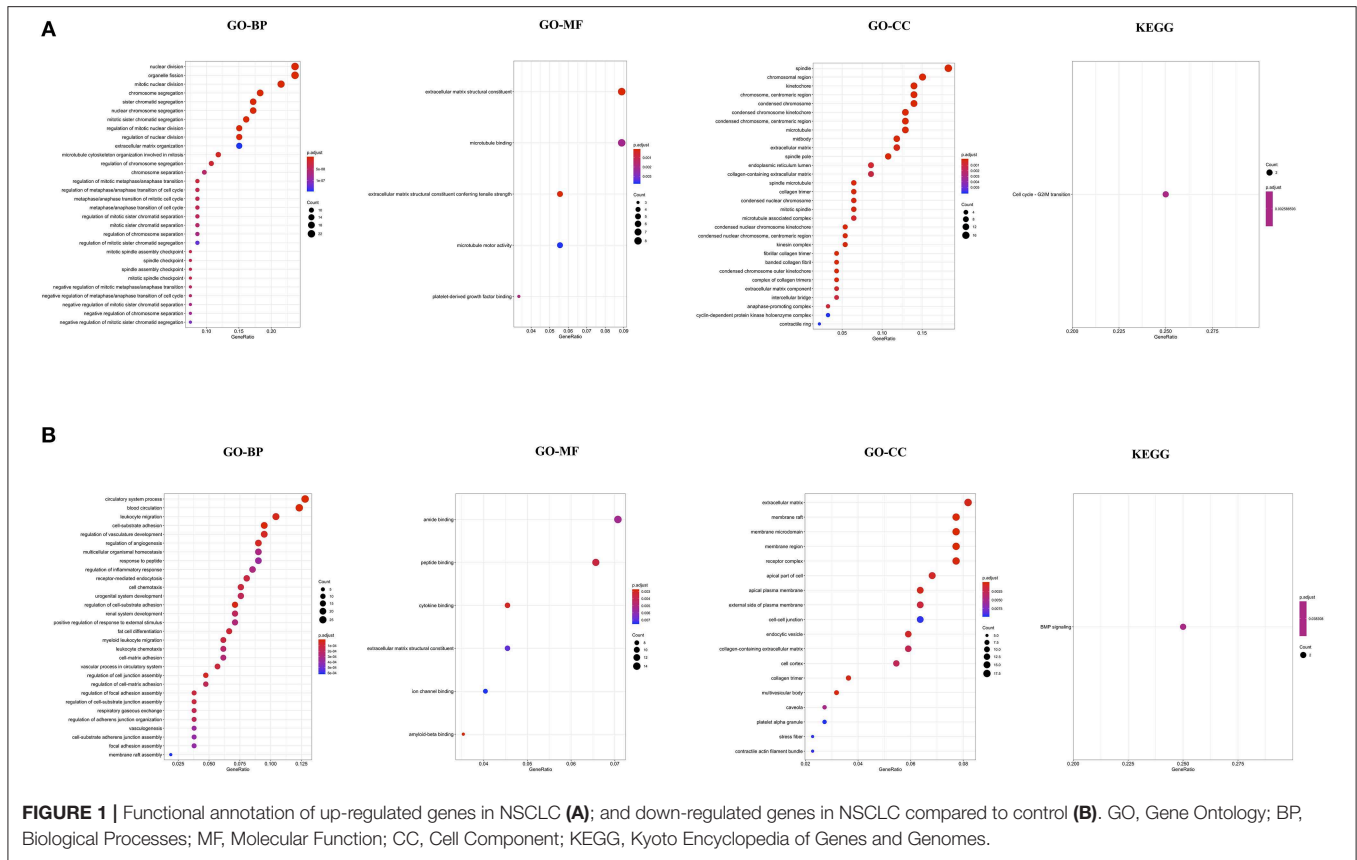
**Hub genes:** The highly connected gene in the network is called hub gene. The node CDC20 has the highest degree of connectivity [35] in the *NSCLC network*. Other top-five hub nodes with their degree of connectivity are BUB1 [33], CDK1 [33], UBE2C [32], CCNB1 [31], and CCNB2 [31] (Table S4). It is interesting to note that all 26 nodes of *NSCLC network* act as intramodular hubs of *Cluster 1*. Therefore, we considered all genes in *Cluster 1* as hub genes as their degree of connectivity are more than 21 (Figure S3; Table S4).

**Betweenness centrality of Node:** The node RHOJ has the highest betweenness centrality of 1, which connects DLC1, ARHGEF26, and ARHGAP6 (Figure 2; Table S4). Node Interleukin-6 (IL6) has second-highest betweenness centrality of 0.611 in the *NSCLC network*, which connects 6 proteins across three sub-networks: *Cluster 1* (UBE2C); *Cluster 2* (CXCL2); and *Cluster 5* (SPP1, CP, GOLM1, CHRDL1) (Figure 2; Table S4).

**Top DEGs in Clusters:** It was found that the highest up-regulated gene are TOP2A (*Cluster 1*), CXCL13 (*Cluster 2*), COL11A1 (*Cluster 3*), and SPP1 (*Cluster 5*); while highest down-regulated genes are PPBP (*Cluster 2*), COL6A6 (*Cluster 3*), EDNRB (*Cluster 4*), and IL6 (*Cluster 5*).

In order to understand the functional relevance, these clusters were further analyzed using GO and pathways enrichments. The *Cluster 1* consist of 26 up-regulated gene in *NSCLC network* (Table 2). Functional annotation indicates that: (a) *Cluster 1* is significantly associated with nuclear division, spindle, microtubule binding, and protein serine/threonine kinase activity (Figure S4; Table S5); (b) *Cluster 2* is significantly associated with leukocyte migration, cell chemotaxis, G





**FIGURE 1 |** Functional annotation of up-regulated genes in NSCLC (A); and down-regulated genes in NSCLC compared to control (B). GO, Gene Ontology; BP, Biological Processes; MF, Molecular Function; CC, Cell Component; KEGG, Kyoto Encyclopedia of Genes and Genomes.

protein-coupled receptor binding, and chemokine activity; (c) *Cluster 3* is significantly associated with extracellular matrix organization and collagen trimer; (d) *Cluster 4* is significantly associated with G protein-coupled receptor signaling pathway via cyclic nucleotide second messenger; (e) *Cluster 5* is significantly associated with post-translational protein modification, cytokine activity (Figure S4).

### Upstream Regulator of Cluster

Transcription factors play crucial roles in initiation, progression, and metastasis of cancer. However, the role of TFs and their downstream target genes and their regulatory mechanisms in the development of NSCLC remains largely unknown. Therefore, each MCODE cluster was analyzed to identify the potential upstream TF regulators using the iRegulone tool.

**Cluster 1:** Our analysis showed that almost all 26 up-regulated genes are under the control of five TFs: FOXM1, MYBL2, TFDP1, E2F4, and SIN3A (Figure 3A). However, only FOXM1 and MYBL2 are up-regulated gene showing log<sub>2</sub>FC > 1, while, TFDP1 and E2F4 show slight up-regulated while SIN3A show slight down-regulated in our list of DEGs of NSCLC.

**Cluster 2:** The cluster 2 contains 10 genes, which is under the regulation of 24 TFs, however, only four TFs showed |log<sub>2</sub>FC| > 1 (TFAP2A and TFAP2C up-regulated; GATA2 and FOS down-regulated) in our list of DEGs (Figure 3B).

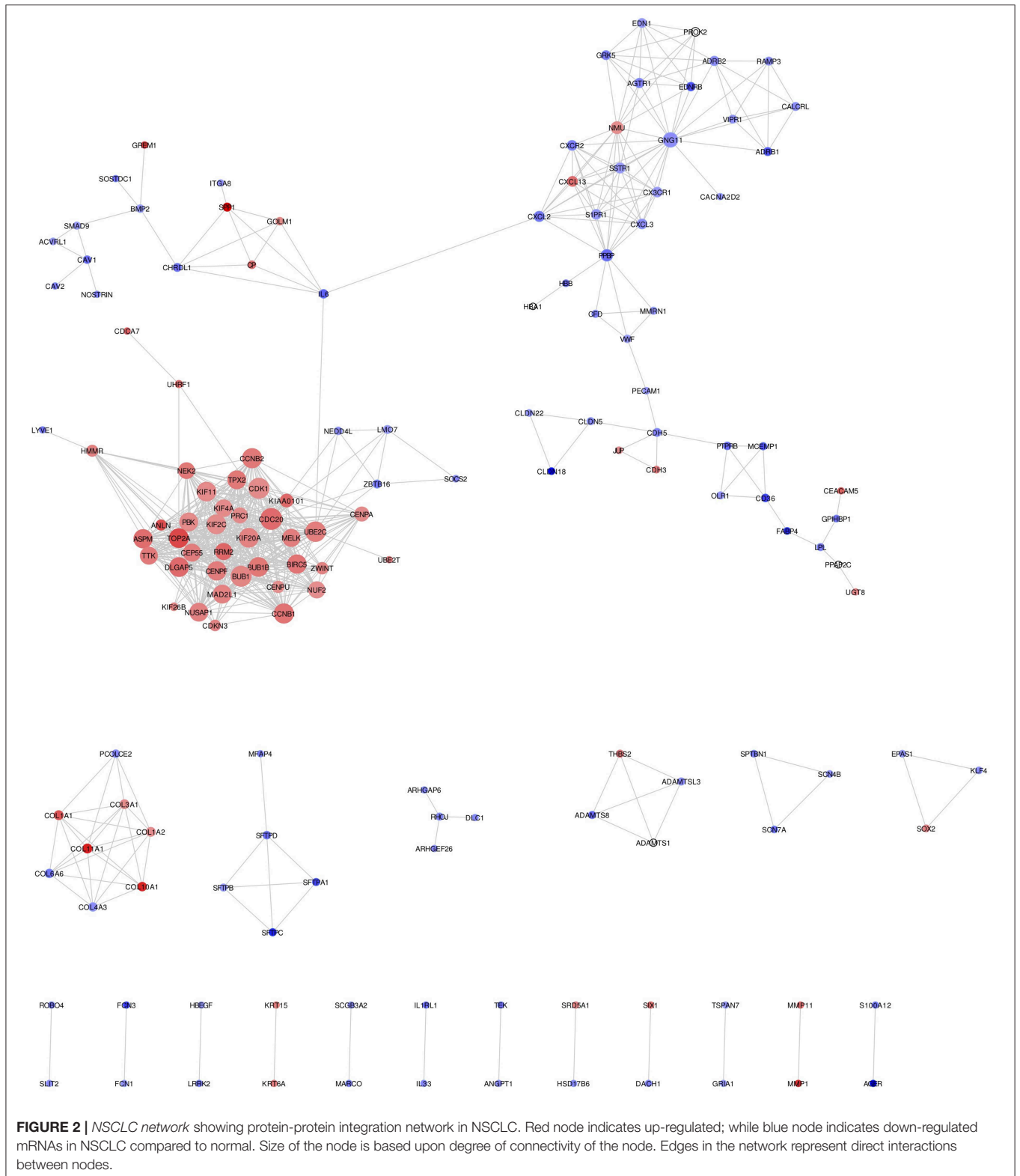
**Cluster 3:** The cluster 3 contains 7 genes, which is under the regulation of 7 TFs, however, all of the TF showed |log<sub>2</sub>FC| < 1 in our list of DEGs (Figure 3C).

**Cluster 4:** The cluster 4 contains 10 genes, which are under the regulation of 20 TFs, however, only GATA6, EBF1, and JUN showed log<sub>2</sub>FC < -1 in our list of DEGs (Figure 3D). **Cluster 5:** Cluster 5 contains 5 genes, which is under the regulation of 22 TFs, however, only six TFs showed |log<sub>2</sub>FC| > 1 (TFAP2A and TFAP2C up-regulated; JUN, FOXA2, JUNB, and FOS down-regulated) in our list of DEGs (Figure 3E). Cluster 1 contains all up-regulated genes, and Cluster 4 contains all down-regulated genes, however, the rest of the cluster contains both up-and down-regulated genes. Venn diagram showing that few TFs are commonly regulating more than one cluster (Figure 3F).

### Validation of Upstream Regulator of Cluster

A study found that the expression level of genes and their TFs are highly correlated in spite of cell diversity; while the expression level of randomly selected genes and TFs show very weak correlation (37). Therefore, TFs interacting with its potential target gene in clusters of NSCLC network were analyzed for their expression correlation.

**Cluster 1:** Using Pearson correlation coefficients, all genes in Cluster 1 are showing significantly highly positive correlation with upstream TFs FOXM1, and MYBL2 in NSCLC (Figure 4A; Figure S5). As revealed in the figures, their gene expression is induced in NSCLC compared to the normal sample. Top five highly correlated expressed genes with a FOXM1 are (a) CCNB2 (R = 0.74); (b) KIF4A (R = 0.73); (c) ASPM (R = 0.72); (d) KIF11 (R = 0.70); and (e) BUB1 (R = 0.69). Top five highly correlated



expressed genes with TF MYBL2 are; (a) BUB1 ( $R = 0.72$ ); (b) KIF4A ( $R = 0.71$ ); (c) KIF2C ( $R = 0.70$ ); (d) KIF11 ( $R = 0.68$ ); and (e) NEK2 ( $R = 0.67$ ).

**Cluster 2:** The highest correlation of  $R = 0.53$  was observed between FOS and CXCL2 (Figure S6). However, our data showed these both genes are down-regulated in NSCLC.

**TABLE 2** | List of 15 highest score clusters identified from NSCLC Network by MCODE.

Cluster	Score (Density*#Nodes)	# Nodes	# Edges	Node IDs
1	25.2	26	315	CEP55, KIF20A, DLGAP5, KIF4A, RRM2, KIF11, NUF2, NUSAP1, UBE2C, CCNB1, BUB1, NEK2, TPX2, BUB1B, TOP2A, MAD2L1, PBK, CCNB2, MELK, CDC20, BIRC5, CDK1, CENPF, ASPM, KIF2C, TTK
2	10	10	45	GNG11, CXCL13, NMU, CXCL2, S1PR1, CX3CR1, CXCL3, PPBP, SSTR1, CXCR2
3	7	7	21	COL3A1, COL11A1, COL6A6, COL1A1, COL4A3, COL1A2, COL10A1
4	5.111	10	23	PROK2, VIPR1, ADRB2, ADRB1, EDNRB, EDN1, RAMP3, CALCRL, GRK5, AGTR1
5	5	5	10	GOLM1, SPP1, IL6, CHRDL1, CP
6	4	4	6	PTPRB, MCEMP1, CD36, OLR1
7	4	4	6	THBS2, ADAMTSL3, ADAMTS8, ADAMTS1
8	4	4	6	SFTPA1, SFTPC, SFTPB, SFTPD
9	3	3	3	ZWINT, CENPA, CENPU
10	3	3	3	ZBTB16, NEDD4L, LMO7
11	3	3	3	VWF, MMRN1, CFD
12	3	3	3	SCN7A, SPTBN1, SCN4B
13	3	3	3	SOX2, KLF4, EPAS1
14	3	3	3	ACVRL1, CAV1, SMAD9
15	2.8	6	7	CLDN18, CLDN5, CDH5, CLDN22, CDH3, JUP

**Cluster 3:** It showed the highest correlation of  $R = 0.36$  between TCF12 and COL1A2 (**Figure S7**). However, our data showed COL1A2 is up-regulated, while TCF12 is slightly down-regulated in NSCLC.

**Cluster 4:** The highest correlation of  $R = 0.34$  was observed between GATA6 and RAMP3 (**Figure S8**). Our data showed these two genes are down-regulated in NSCLC.

**Cluster 5:** Cluster 5 showed the highest correlation of  $R = 0.33$  between FOXA2 and GOLM1 (**Figure S9**). However, our data showed GOLM1 is up-regulated, while FOXA2 is down-regulated in NSCLC.

## Gene Expression-Based Survival Analysis in NSCLC by Kaplan-Meier Plot

The topologically significant genes in the global NSCLC network, genes in MCODE clusters, and upstream regulator TFs (showing  $|\log_2FC| > 1$ ) were analyzed for association with OS in NSCLC using Kaplan-Meier plots. Kaplan-Meier plots of each cluster and their associated TFs are presented as follow: **Cluster 1** in **Figure 4B** and **Figure S10**; **Cluster 2** in **Figure S11**; **Cluster 3** in **Figure S12**; **Cluster 4** in **Figure S13**; and **Cluster 5** in **Figure S14**. Kaplan-Meier plots showed that high expression of all the up-regulated genes of **Cluster 1** make worse the OS [HR > 1], while high expression of down-regulated gene SIN3A makes better the OS [HR < 1] in NSCLC (**Figure 4B**; **Figure S10**). Kaplan-Meier plots of the gene of other clusters showed very much similar patterns that high expression of up-regulated genes make worse the OS, while high expression of down-regulated genes make better the OS in NSCLC (**Figures S11–S14**).

## Extension of Cluster 1 With miRNA

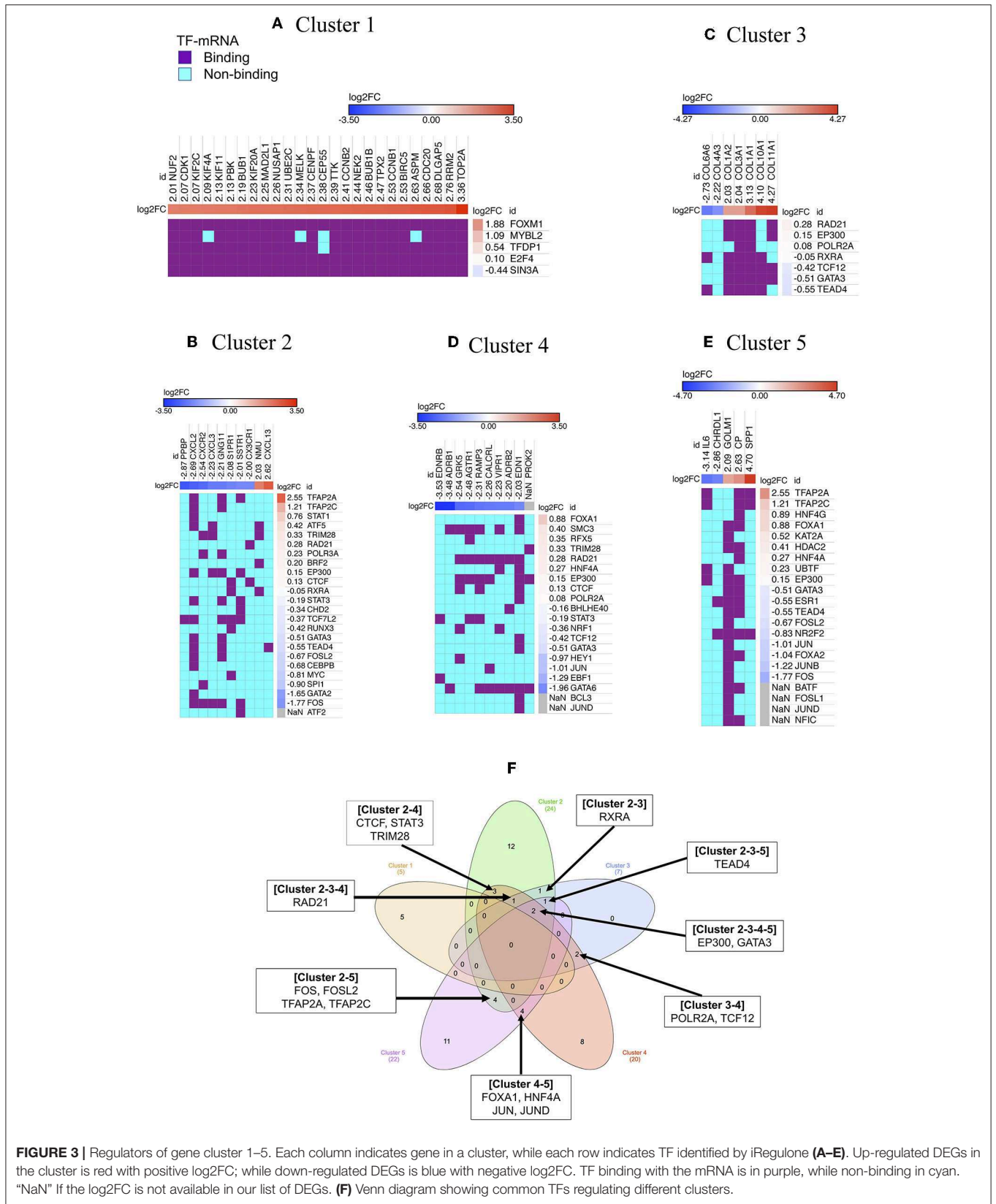
Our analysis identified 30 up-regulated and 70 down-regulated miRNAs targeting 25 genes of **Cluster 1** and associated five TFs. These data were used to generate miRNA network of **Cluster 1** consisting of 130 nodes and 218 interactions (**Figure 5**). Our

analysis found none of the miRNA is targeting NUSAP1 gene of **Cluster 1**. The top five genes targeted by highest number of down-regulated miRNAs in NSCLC are: (a) RRM2 targeted by 17 miRNAs; (b) BIRC5 targeted by 14 miRNAs; (c) CEP55 targeted by 12 miRNAs; (d) KIF2C targeted by 11 miRNAs; (e) CDK1 targeted by 9 miRNAs (**Table S6**). Interestingly, it was found that the expression of TFs, regulators of **Cluster 1** genes, are also under the control of miRNAs as following: (a) FOXM1 is targeted by 10 miRNA (9 down- and 1 up-regulated); (b) MYBL2 is targeted by 6 miRNAs (5 down- and 1 up-regulated); (c) TFDP1 is targeted by 7 miRNAs (6 down- and 1 up-regulated); (d) E2F4 is targeted by 5 miRNAs (4 down- and 1 up-regulated); (e) SIN3A is targeted by 6 miRNAs (4 down- and 2 up-regulated). The complete list of gene and its associated miRNAs are provided in **Table S6**.

The miRNA targeting highest number of genes are as following: (a) hsa-miR-193b-3p is targeting 11 genes (ASPM, BUB1, BUB1B, CDC20, CDK1, KIF11, MELK, RRM2, TOP2A, TPX2, UBE2C); (b) hsa-miR-215-5p is targeting 10 genes (ASPM, BUB1B, CDC20, CENPE, CEP55, DLGAP5, KIF20A, MAD2L1, NUF2, TTK); (c) hsa-miR-186-5p is targeting 7 genes (BUB1, DLGAP5, FOXM1, KIF11, NEK2, RRM2, TOP2A); (d) hsa-miR-16-5p is targeting 7 genes (BIRC5, CDC20, CDK1, CENPE, CEP55, KIF2C, UBE2C); and (e) hsa-miR-30a-5p is targeting 5 genes (CDC20, KIF11, MYBL2, RRM2, TFDP1). The complete list of miRNAs and their targets are provided in the **Table S7**. Interestingly, hsa-miR-193b-3p, hsa-miR-215-5p, hsa-miR-186-5p, hsa-miR-16-5 and hsa-miR-30a-5p are down-regulated and their targets are up-regulated in NSCLC compared to control indicating their role in the development of NSCLC.

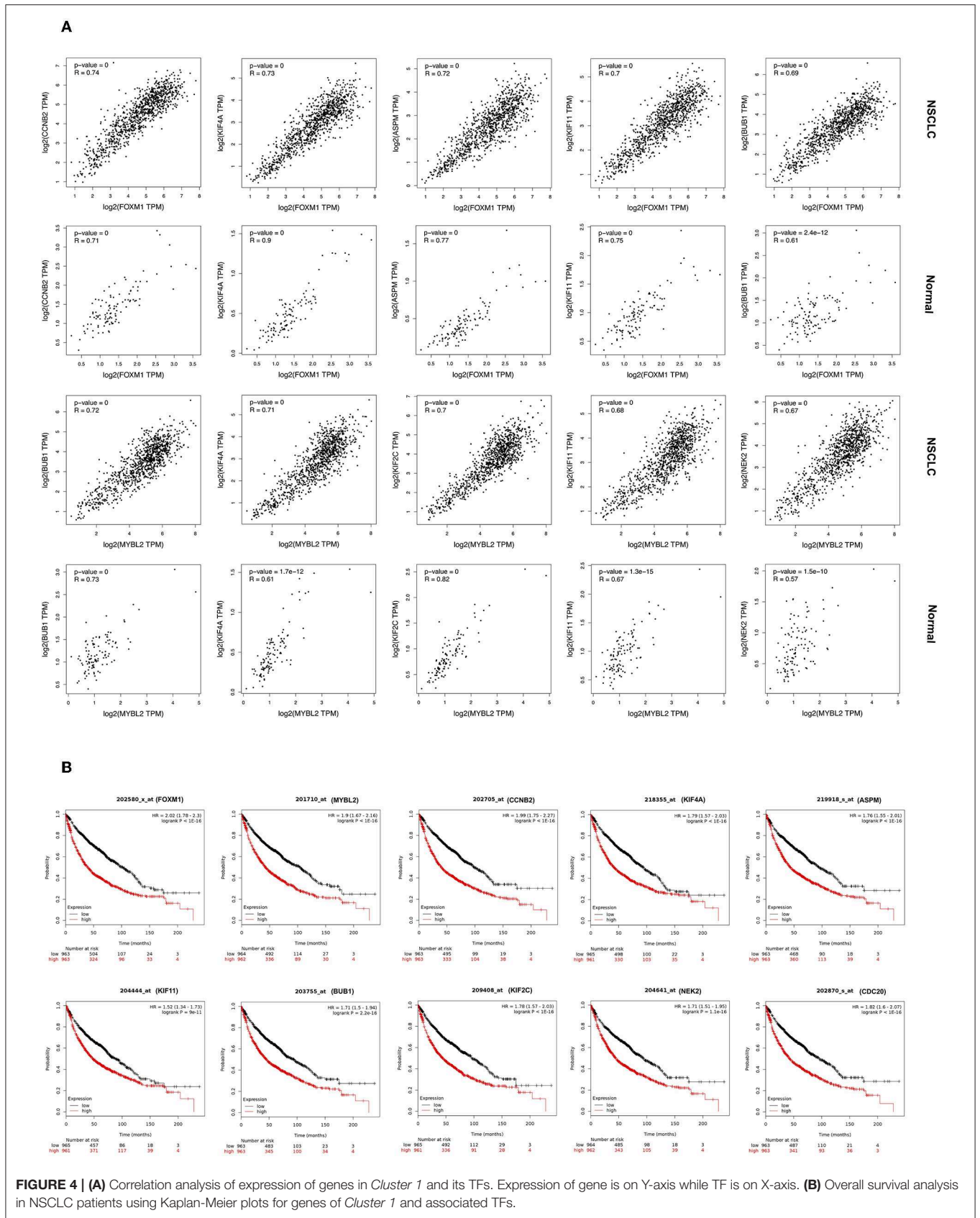
## Mutational Signatures in NSCLC

Analysis of mutational signatures in 31 genes (**Cluster 1** and associated TFs) in NSCLC studies showed that queried genes are altered in 1966 (37%) out of 5279 samples across TCGA datasets

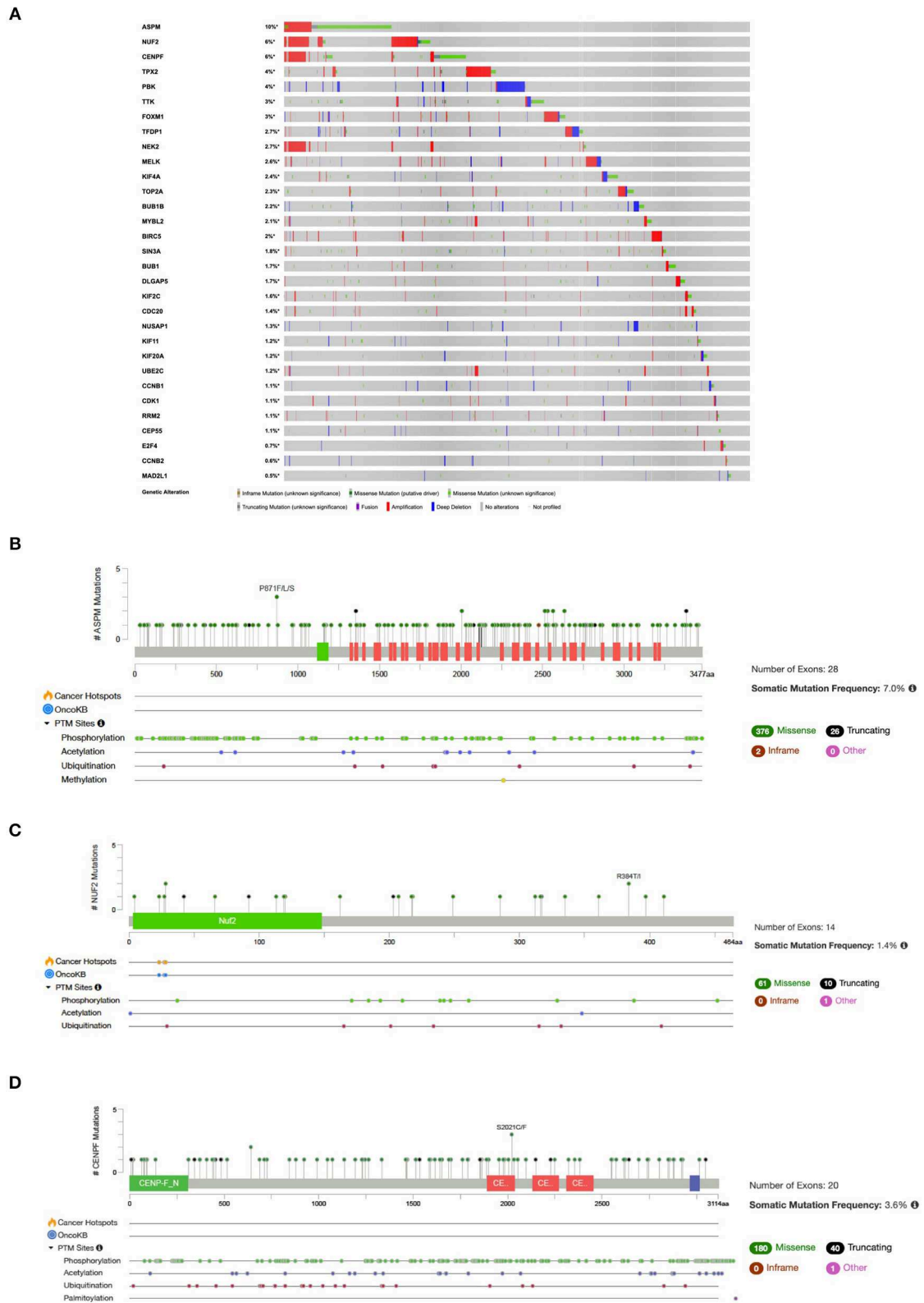


**FIGURE 3 |** Regulators of gene cluster 1–5. Each column indicates gene in a cluster, while each row indicates TF identified by iRegulone (A–E). Up-regulated DEGs in the cluster is red with positive log2FC; while down-regulated DEGs is blue with negative log2FC. TF binding with the mRNA is in purple, while non-binding in cyan. “NaN” if the log2FC is not available in our list of DEGs. (F) Venn diagram showing common TFs regulating different clusters.









**FIGURE 6 | (A)** OncoPrint of genes in *Cluster 1* and associated TFs alteration in NSCLC. Lollipop plot with distribution of mutations in NSCLC across protein domains of **(B)** ASPM; **(C)** NUF2; and **(D)** CENPF.

In the current study, a meta-analysis of large gene expression samples (131 NSCLC and 130 control) identified 346 DEGs (97 up-regulated and 249 down-regulated) in NSCLC compared to adjacent non-tumor lung tissues. After integrating the PPI data into DEGs, the *NSCLC network* was created and analyzed to understand the dysregulated sub-networks and pathways in NSCLC. Furthermore, the sub-networks were studied to identify a “driver-network” and its upstream regulators by integrating the data of TFs and miRNAs. The topologically important gene in *NSCLC network*, the “driver-network,” and its upstream regulators could be candidate genes for biomarkers and therapeutic target for NSCLC. However, selecting candidate genes for biomarker and therapeutic target requires their overexpression or underexpression have a deep connection at the molecular level for the initiation and progression of tumorigenesis. Therefore, the literature and databases mining was performed on the selected candidate genes and regulatory sub-network to understand their molecular mechanism in NSCLC.

## DEGs, Functional Annotation and Pathway Enrichment

Our study found that SPP1 is the most up-regulated gene and its high expression is associated with reducing OS (Table 1; Figure S14) and thus, support the previous finding showed enhanced expression of SPP1 in several types of tumors including NSCLC (42). Network analysis showed SPP1 is the member of *Cluster 5* and under the regulation of TFs TFAP2A and TFAP2C (Figure 3E). However, in spite of overexpression of both TFs in NSCLC, our study found no positive correlation between the expression of these TFs and SPP1 and therefore need further investigation (Figure S9). SPP1 binds to CD44 and integrin receptor in the lung cancer cell and activates the FAK/PI3K/AKT pathway which induces the secretion of vascular endothelial growth factor (VEGF) resulting in increased cell survival, cell proliferation and tumor metastasis (42). Silencing the expression of SPP1 using siRNA decreased the NSCLC tumor volume and weight in mice demonstrated it as a promising therapeutic target (43). Furthermore, our analysis showed that AGER is the highly down-regulated gene in NSCLC compared to normal tissue. AGER is a multi-ligand receptor that binds various ligands derived from a damaged cell and its up-regulation at both mRNA and protein level is associated with the majority of cancers including gastric, breast, hepatocellular, colorectal carcinoma (44, 45). However, unlike other cancers, AGER is down-regulated in NSCLC and also supported by the previous finding suggested its role as a tumor suppressor in lung cancer (46). S100A12 is a small protein express by neutrophil granulocytes and binds with AGER receptor, which induces the production of proinflammatory cytokines (47). AGER and S100A12 are interacting in *NSCLC network* and both are down-regulated suggested their role in reducing the inflammation to escape from the immune response in NSCLC (Figure 2) (47). A previous study also supports our finding that SPP1 and AGER are highly up-regulated and down-regulated, respectively, in NSCLC (48). A pan-cancer analysis

of pediatric leukemias and solid tumors identified 142 and 82 driver genes (40). Comparing with pan-cancer driver genes, our study found that SIX1 was up-regulated, while TAL1 and ID4 were down-regulated in NSCLC (40). The pan-cancer study reported mutation in SIX1, TAL1, and ID4 were present in the Wilms Tumors, T-lineage acute lymphoblastic leukemias, and B-lineage acute lymphoblastic leukemias, respectively (40). Taken together the functional annotation and pathway enrichment analysis of DEGs indicated that up-regulated genes could lead to enhance tumor cell proliferation, while down-regulated genes decreased in immune cell migration in NSCLC, which are vital for uncontrolled cell division and survival in cancer, and to escape from the proper immune response (Figure 1).

## Topology of NSCLC Network

Topological properties of *NSCLC network* identified several other key proteins. Hub gene plays a key role in the proper maintaining the architecture of the biological network (49, 50). The study found that the intramodular hubs are significantly related to cell proliferation and survival time in cancer (49, 51). All of the protein of *Cluster 1* genes are highly interacting and therefore act as hub genes in the *NSCLC network* (Table S4). Top six hub genes with more than 30 degrees of connectivity are CDC20 [i.e. 35], BUB1 [33], CDK1 [33], UBE2C [32], CCNB1 [31], and CCNB2 [31] (Table S4). CDC20 regulates cell division through activating the anaphase-promoting complex/cyclosome (APC/C), which begins chromatid separation to enter into anaphase. Overexpression of CDC20 is reported in various cancer including breast cancer, cervical cancer, urinary bladder cancer, and associated with poor prognosis of ovarian tumors (52). It was reported that overexpression of CDC20 is associated with poor prognosis in NSCLC, which support our findings (Figure 4B) (8). Therefore, various studies considered CDC20 as a therapeutic target for cancer treatment (52).

The ubiquitin-conjugating enzyme E2 (UBE2C) is a member of the APC/C complex and promotes the degradation of various target proteins required for cell cycle progression. The aberrantly high expression of UBE2C was reported in various cancers. It was experimentally showed that the TF FOXM1 binds to the promoter region of UBE2C and activates its high expression in esophageal squamous cell carcinoma, which supports our finding that the FOXM1 as an upstream regulator of UBE2C of *Cluster 1* (Figure 3A) (53). CCNB1 interact with CDK1 to form a complex that phosphorylate their substrates and promotes G2/M transition in the cell cycle. *Cluster 1* contains various protein including CCNB1, which are degraded by APC/C E3 ubiquitin ligase complex. Overexpression of CCNB1 resulting in cell proliferation and was reported in various cancers including NSCLC (9). Inhibiting the expression of CCNB1 using siRNAs promotes apoptosis in colorectal cancer cells (54). BUB1 is component of spindle checkpoint for proper chromosome segregation and its up-regulation was reported in human prostate cancer (55).

Node with the highest betweenness centrality controls the flow of information between two nodes and therefore, could be crucial protein in signaling network and potential drug target to stop the flow of communication in a disease state. The node



RHOJ has the highest betweenness centrality and interacts with DLC1, ARHGEF26, and ARHGAP6 and they all are down-regulated in NSCLC (Figure 2; Table S4). A study found that the encoded protein of RHOJ is activated by VEGF and may regulate angiogenesis. DLC1 encodes a Rho GTPase-activating protein that functions as a tumor suppressor and down-regulated in more than 95% of NSCLC and other cancers (56). A previous study supports the role of DLC1 as an inducer of apoptosis in NSCLC (57) and as a metastasis suppressor in breast cancer cells (58). ARHGEF26 function as guanine nucleotide exchange factors (GEFs), which catalyze the activation of RHOA by displacing GDP (from inactive enzyme) with GTP (active enzyme). ARHGAP6 involves in regulating the actin polymerization at the cell plasma membrane. Our network analysis indicates that Rho family proteins forming a complex of RHOJ- DLC1-ARHGEF26-ARHGAP6, which need detail study in NSCLC. Node IL6 has second highest betweenness centrality of 0.611 in the NSCLC network, which connects 6 proteins across three sub-networks. IL-6 is cytokines secreted during inflammation and chronic disease like cancer. It binds with interleukin-6 receptor alpha (IL-6R $\alpha$ ) present on the surface of T-cell, NK cell, B-cell and activates them. IL6 and its interacting partners CXCL2 and CHRDL1 are down-regulated; while other interacting partners UBE2C, SPP1, CP, GOLM1 are up-regulated in the NSCLC network. CXCL2 gene encodes secreted proteins and plays an important role in inflammation and immunoregulation. A study found CXCL2 role in the resistance of anti-cancer drug, anlotinib, in NSCLC (59). The overexpression of GOLM1 is reported in prostate cancer (60) and lung adenocarcinoma (61). Because of high betweenness centrality of down-regulated IL6, it may be playing a pivotal role in the down-regulation of the inflammatory response in NSCLC.

### Driver-Network and Upstream Regulators

Further analysis of NSCLC network identified the biologically informative 15 local cluster networks. Among them, the highest scoring Cluster 1 identified as local “driver-network” having 26 overexpressed gene and their upstream regulators FOXM1 (log2FC = 1.88), MYBL2 (log2FC = 1.09), TFDP1 (log2FC = 0.54), E2F4 (log2FC = 0.10), and SIN3A (log2FC = -0.44) (Figure 3A). The “driver-network” is collectively associated with cell proliferation (Cluster 1 in Figure S4). Interestingly, we observed a strong positive correlation between gene expression of each member of “driver-network” and its upstream regulators FOXM1 and MYBL2 in NSCLC (Figure 4A; Figure S5). Furthermore, NSCLC patients with their overexpression had significantly worse OS (Figure 4B; Figure S10).

The previous study found that MuvB core proteins interact with E2F4-DP1 and p130 or p107 to form a DREAM complex in G0/G1 phase of the cell cycle, which put the cell in quiescence state by globally repressing more than 800 cell cycle genes (62). When cell exit from quiescence state, MuvB core proteins dissociated from p130 and interacts with MYBL2 to form MMB (MYBL2-MuvB) complex. Subsequently, MMB recruits FOXM1 protein to form MMB-FOXM1 complex, which binds to the promoters of several cell cycle genes and activate their expression in G2/M phase responsible for mitosis (63). A study found that high expression MYBL2 gene disrupts the DREAM complex

and increase the MMB complex formation and subsequently triggers the expression of the several target genes driving the cell proliferation in cancer (64). In this way, MMB complex function as opposite of the DREAM complex. A previous study identified the highly confident candidate target genes and regulatory network of DREAM and MMB-FOXM1 complexes involved in the cell cycle (63). Comparing to the study, it was found that all 26 genes of Cluster 1 are the target of DREAM complex (63). The same study support that 24 genes of Cluster 1 (except MELK and PBK) are the target of MMB-FOXM1 complex (63). Furthermore, the TFs FOXM1, MYBL2, SIN3A, and TFDP1 are the target gene of DREAM, but not the MMB-FOXM1 complex. Combining all these findings, our study indicates that most of the genes of Cluster 1 are the common target for both DREAM and MMB-FOXM1 complexes. However, overexpression of MYBL2 and FOXM1 could disrupt the DREAM complex and enhance the formation of MMB-FOXM1 complex resulting high expression of cell cycle genes in Cluster 1 which consequences uncontrolled cell proliferation and ultimately NSCLC.

FOXM1 is a member of the Forkhead box family of transcriptional factor that expresses in actively dividing cells. Several studies reported the overexpression of FOXM1 stimulates the proliferation of tumor cells during the progression of NSCLC and other types of cancers and also associated with poor overall survival (13–15, 65). A study found FOXM1 overexpressed NSCLC associated with resistance of cisplatin-based chemotherapy, and its inhibition using thiothrepton or siRNA reversed the drug resistance resulted in inhibition of cell proliferation and induce cell death (14). Silencing of FOXM1 expression by siRNA in A549 lung adenocarcinoma cells resulted in significant reduction in cell cycle-promoting cyclin A2 and cyclin B1 genes, as well as DNA replication and mitosis (13). DREAM complex directly represses the transcription of TOP2A, which encode DNA topoisomerase to relief the torsional stress during DNA transcription and replication (62). Furthermore, the study showed that depleting FOXM1 expression decrease the TOP2A mRNA and protein level in A549 human lung adenocarcinoma cells (66). Experimental studies showed the FOXM1 protein directly bind to the promoter region of TOP2A mRNA (66). Our study showed overexpression of TOP2A (log2FC = 3.36) and its upstream regulator FOXM1 indicates that both genes are the promising target for anti-cancer therapy for NSCLC (67). A pan-cancer study found FOXM1 is overexpressed across all studied 32 TCGA cancer types including NSCLC compared to normal tissues (68). A pan-cancer analysis revealed FOXM1 regulatory network as a top predictor of poor prognosis (69). We found all the genes of Cluster 1 except NUSAP1, PBK, and CDK1 are present in the pan-cancer network associated with mitotic cell cycle and adverse prognostic genes [see Figure 2d of (69)]. In addition, the pan-cancer network contains the TFs FOXM1 and MYBL2 which support our finding.

The MYBL2 is phosphorylated by cyclin A/cyclin-dependent kinase 2 during the S-phase of the cell cycle and activate the cell division (12). Overexpression of MYBL2 is associated with poor patient survival in various cancers patient including NSCLC (12, 70). A previous study showed experimentally that several genes including KIF20A, KIF4A, NUSAP1, CCNB1, TOP2A,

CDK1, CENPF, and KIF2C of *Cluster 1* are transactivated by MYBL2 (12). Abnormal expression of E2F4 and its mutations are reported in several cancers including NSCLC (16, 71, 72).

The genes of *Cluster 1* are mainly involving in the proliferation of cell division. Our study supports the previous finding that overexpression of *Cluster 1* genes in several cancer including NSCLC and associated with poor overall survival such as: KIF2C (73), KIF4A (11, 74), and KIF11 (75) which are kinesin family members of motor proteins regulating the cell mitosis through faithful chromosome condensation and segregation (76). Furthermore, it was reported that silencing of their expression using siRNAs inhibit the cancer cell growth (11, 73). Several inhibitors of KIF11 such as Monastrol, Ispinesib, and Dimethylenastron have been developed and are in clinically used to inhibit cell proliferation and induce apoptosis to treat numerous cancers (77). The high expression of ASPM responsible for mitotic spindle formation (10, 78). NEK2 (79, 80) which triggers centrosome separation are reported in NSCLC. Interestingly, another study also supports that E2F4 and FOXM1 bind to the promoter of NEK2 gene (80). NUF2 component of NDC80 Kinetochore complex regulating the chromosome segregation was overexpressed in NSCLC (81). Suppressing the expression of NUF2 inhibits tumor growth and also stimulates cell apoptosis (81). CENPF (82, 83) which is associated with the centromere-kinetochore complex and requires for chromosome segregation during mitosis.

Therefore, the “driver-network” and the predicted TFs MYBL2 and FOXM1 give more insight about the initiation and progression of NSCLC and also could be therapeutic target genes. However, a further biochemical study is required to understand the effect on cell proliferation in NSCLC by using functional siRNAs targeting a combination of TFs FOXM1, and MYBL2, and their downstream genes of “driver-network” (84).

Our study found that the “driver-network” is not only under the regulation of TFs, but also under the regulation of miRNAs. This study showed that 9 miRNAs (hsa-miR-134-5p, hsa-miR-149-5p, hsa-miR-186-5p, hsa-miR-194-5p, hsa-miR-204-5p, hsa-miR-26b-5p, hsa-miR-320a, hsa-miR-370-3p, hsa-miR-630) targeting FOXM1 are down-regulated in NSCLC. In addition, these miRNAs are also targeting 11 other genes (ASPM, BIRC5, BUB1, DLGAP5, KIF11, KIF4A, MAD2L1, NEK2, RRM2, TOP2A, and TTK), which means these are common targets for miRNAs and FOXM1 (Table S7). Previous studies found the down-regulation of has-miR-134-5p and hsa-miR-149-5p were contributing epithelial-to-mesenchymal transition (EMT), a key process of cancer metastasis, in NSCLC (85, 86). These studies also demonstrated that has-miR-134-5p and hsa-miR-149-5p act as tumor suppressors by directly binds to the 3'UTR of FOXM1 and inhibiting its expression and the EMT in NSCLC (85, 86). Accumulating evidence indicate that down-regulation of these miRNAs eliminate their suppressive effect resulting overexpression of FOXM1 and its 11 downstream target genes. The decrease expression of other miRNAs in NSCLC and their role in cell proliferation and EMT has been demonstrated in several studies including hsa-miR-194-5p (87), hsa-miR-204-5p (88), hsa-miR-26b-5p (89), hsa-miR-320a (90), hsa-miR-370-3p (91) and hsa-miR-630 (92). MYBL2 and TFDP1 are targeted by five common miRNAs (hsa-miR-30a-5p, hsa-miR-30b-5p,

hsa-miR-30c-5p, hsa-miR-30d-5p, hsa-miR-30e-5p), though they belong from same miRNA family. A previous study showed the down-regulation of hsa-miR-30a-5p which directly targeting MYBL2 mRNA in NSCLC (93). Interestingly, TCGA NSCLC dataset showed higher mutation rates in the genes of “driver-network” as well as its upstream regulators FOXM1 and MYBL1 in the NSCLC (Figure 6).

Previous bioinformatics studies were mainly focused on the analysis of gene expression data to identify the DEGs, their function enrichment, the interacting hub genes in NSCLC. Ni *et al.* identified five up-regulated hub genes (TOP2A, CCNB1, CCNA2, UBE2C, and KIF20A) in NSCLC (41). Huang *et al.* identified five up-regulated hub genes (CDC20, CENPF, KIF2C, BUB1, and ZWINT) in NSCLC (94). Another study found 16 hub genes (TEK, ANGPT1, MMP9, VWF, CDH5, EDN1, ESAM, CCNE1, CDC45, PRC1, CCNB2, AURKA, MELK, CDC20, TOP2A, and PTTG1) in NSCLC (95). However, our study has following advantages compared to previous studies: (a) Current study is based upon a large dataset of NSCLC obtained from different GEO microarray dataset; (b) Identified various common DEGs detected by previous studies; (c) The DEGs were integrated with PPI, TFs and miRNAs to understand the regulatory mechanism of NSCLC initiation and progression; (d) Finally, we have identified a “driver-network” consist of 26 up-regulated hub genes and their upstream regulators (FOXM1, MYBL2, and miRNAs) involved in the proliferation of NSCLC and could serve as diagnostic and therapeutic targets to treat NSCLC.

Our study suggested that a gene could be functionally important even at a small level of overexpression such as FOXM1 and MYBL2 if they act as upstream regulators of genes of interacting proteins involved in an important biological process (20). However, our study has following limitations: (a) All the findings are based upon computational analysis using integrated data of gene expression, PPI, TFs, and miRNAs; (b) Our study did not integrate the data of gene mutations and gene copy number variations, which could abolish the cis- and trans-regulatory elements of a gene resulting aberrant gene expression and cancer (96–98); and (c) Finally, our study lacks the experimental validation and therefore need further experimental testing.

## CONCLUSION

In this study, potential biomarkers and therapeutic targets has been identified for NSCLC using systems bioinformatics approach on the public gene expression data. All the genes in “driver-network” (*Cluster 1*) and its upstream regulators, FOXM1 and MYBL2, which collectively overexpressed and involve in the cell proliferation, and cell division are particularly promising for further study. Furthermore, we identified several tumor suppressor miRNAs and their interacting target genes in the “driver-network.” Targeting two or more genes of the “driver-network” may be synergistic and more effective therapy against NSCLC. In our study, correlation expression, OS, and gene mutations dataset with strong statistical support were used to validate our finding. However, the biochemical study on the potential biomarkers and therapeutic targets are necessary for further validation on clinical samples.

## DATA AVAILABILITY STATEMENT

This study was conducted on publicly available data on Gene Expression Omnibus (GEO <https://www.ncbi.nlm.nih.gov/geo/>) with accession number: GSE27262, GSE18842, and GSE19804.

## AUTHOR CONTRIBUTIONS

FA conceptualized the whole study, designed and performed all the Bioinformatics analysis, interpreted the results, wrote, and revised the manuscript.

## REFERENCES

- Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, Mok T, et al. Non-small-cell lung cancer. *Nat Rev Dis Primers*. (2015) 1:15009. doi: 10.1038/nrdp.2015.48
- Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. (2014) 511:543–50. doi: 10.1038/nature13385
- Cancer Genome Atlas Research N. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. (2012) 489:519–25. doi: 10.1038/nature11404
- Paez JG, Janne PA, Lee JC, Tracy S, Greulich H, Gabriel S, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. *Science*. (2004) 304:1497–500. doi: 10.1126/science.1099314
- Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, et al. EGF receptor gene mutations are common in lung cancers from “never smokers” and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci USA*. (2004) 101:13306–11. doi: 10.1073/pnas.0405220101
- Vavala T, Novello S. Alectinib in the treatment of ALK-positive non-small cell lung cancer: an update on its properties, efficacy, safety and place in therapy. *Ther Adv Med Oncol*. (2018) 10:1758835918789364. doi: 10.1177/1758835918789364
- Sanchez-Palencia A, Gomez-Morales M, Gomez-Capilla JA, Pedraza V, Boyero L, Rosell R, et al. Gene expression profiling reveals novel biomarkers in non-small cell lung cancer. *Int J Cancer*. (2011) 129:355–64. doi: 10.1002/ijc.25704
- Kato T, Daigo Y, Aragaki M, Ishikawa K, Sato M, Kaji M. Overexpression of CDC20 predicts poor prognosis in primary non-small cell lung cancer patients. *J Surg Oncol*. (2012) 106:423–30. doi: 10.1002/jso.23109
- Yoshida T, Tanaka S, Mogi A, Shitara Y, Kuwano H. The clinical significance of Cyclin B1 and Wee1 expression in non-small-cell lung cancer. *Ann Oncol*. (2004) 15:252–6. doi: 10.1093/annonc/mdh073
- Jung HM, Choi SJ, Kim JK. Expression profiles of SV40-immortalization-associated genes upregulated in various human cancers. *J Cell Biochem*. (2009) 106:703–13. doi: 10.1002/jcb.22063
- Taniwaki M, Takano A, Ishikawa N, Yasui W, Inai K, Nishimura H, et al. Activation of KIF4A as a prognostic biomarker and therapeutic target for lung cancer. *Clin Cancer Res*. (2007) 13(22 Pt 1):6624–31. doi: 10.1158/1078-0432.CCR-07-1328
- Musa J, Aynaud MM, Mirabeau O, Delattre O, Grunewald TG. MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. *Cell Death Dis*. (2017) 8:e2895. doi: 10.1038/cddis.2017.244
- Kim IM, Ackerson T, Ramakrishna S, Tretiakova M, Wang IC, Kalin TV, et al. The Forkhead Box m1 transcription factor stimulates the proliferation of tumor cells during development of lung cancer. *Cancer Res*. (2006) 66:2153–61. doi: 10.1158/0008-5472.CAN-05-3003
- Wang Y, Wen L, Zhao SH, Ai ZH, Guo JZ, Liu WC. FoxM1 expression is significantly associated with cisplatin-based chemotherapy resistance and poor prognosis in advanced non-small cell lung cancer patients. *Lung Cancer*. (2013) 79:173–9. doi: 10.1016/j.lungcan.2012.10.019
- Yang DK, Son CH, Lee SK, Choi PJ, Lee KE, Roh MS. Forkhead box M1 expression in pulmonary squamous cell carcinoma: correlation with clinicopathologic features and its prognostic significance. *Hum Pathol*. (2009) 40:464–70. doi: 10.1016/j.humpath.2008.10.001
- Bankovic J, Stojisic J, Jovanovic D, Andjelkovic T, Milinkovic V, Ruzdijic S, et al. Identification of genes associated with non-small-cell lung cancer promotion and progression. *Lung Cancer*. (2010) 67:151–9. doi: 10.1016/j.lungcan.2009.04.010
- Ahmed F, Senthil-Kumar M, Lee S, Dai X, Mysore KS, Zhao PX. Comprehensive analysis of small RNA-seq data reveals that combination of miRNA with its isomiRs increase the accuracy of target prediction in *Arabidopsis thaliana*. *RNA Biol*. (2014) 11:1414–29. doi: 10.1080/15476286.2014.996474
- Ahmed F, Ansari HR, Raghava GP. Prediction of guide strand of microRNAs from its sequence and secondary structure. *BMC Bioinformatics*. (2009) 10:105. doi: 10.1186/1471-2105-10-105
- Han Y, Li H. miRNAs as biomarkers and for the early detection of non-small cell lung cancer (NSCLC). *J Thorac Dis*. (2018) 10:3119–31. doi: 10.21037/jtd.2018.05.32
- Dutta B, Pusztai L, Qi Y, Andre F, Lazar V, Bianchini G, et al. A network-based, integrative study to identify core biological pathways that drive breast cancer clinical subtypes. *Br J Cancer*. (2012) 106:1107–16. doi: 10.1038/bjc.2011.584
- Sanchez-Vega F, Mina M, Armenia J, Chatila WK, Luna A, La KC, et al. Oncogenic signaling pathways in the cancer genome atlas. *Cell*. (2018) 173:321–37 e10. doi: 10.1016/j.cell.2018.03.035
- Wei TY, Hsia JY, Chiu SC, Su LJ, Juan CC, Lee YC, et al. Methylosome protein 50 promotes androgen- and estrogen-independent tumorigenesis. *Cell Signal*. (2014) 26:2940–50. doi: 10.1016/j.celsig.2014.09.014
- Wei TY, Juan CC, Hsia JY, Su LJ, Lee YC, Chou HY, et al. Protein arginine methyltransferase 5 is a potential oncoprotein that upregulates G1 cyclins/cyclin-dependent kinases and the phosphoinositide 3-kinase/AKT signaling cascade. *Cancer Sci*. (2012) 103:1640–50. doi: 10.1111/j.1349-7006.2012.02367.x
- Lu TP, Tsai MH, Lee JM, Hsu CP, Chen PC, Lin CW, et al. Identification of a novel biomarker, SEMA5A, for non-small cell lung carcinoma in nonsmoking women. *Cancer Epidemiol Biomarkers Prev*. (2010) 19:2590–7. doi: 10.1158/1055-9965.EPI-10-0332
- Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics*. (2010) 26:2363–7. doi: 10.1093/bioinformatics/btq431
- Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics*. (2012) 28:882–3. doi: 10.1093/bioinformatics/bts034
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. (2015) 43:e47. doi: 10.1093/nar/gkv007
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. (2009) 4:44–57. doi: 10.1038/nprot.2008.211

## ACKNOWLEDGMENTS

The author would like to acknowledge the technical support provided by the Department of Biochemistry, and University of Jeddah Center for Scientific and Medical Research (UJCSMR), University of Jeddah.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.01011/full#supplementary-material>



29. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. (2012) 16:284–7. doi: 10.1089/omi.2011.0118
30. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res*. (2015) 43:D447–52. doi: 10.1093/nar/gku1003
31. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. (2003) 13:2498–504. doi: 10.1101/gr.1239303
32. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. (2003) 4:2. doi: 10.1186/1471-2105-4-2
33. Janky R, Verfaillie A, Imrichova H, Van de Sande B, Standaert L, Christiaens V, et al. iRegulon: from a gene list to a gene regulatory network using large motif and track collections. *PLoS Comput Biol*. (2014) 10:e1003731. doi: 10.1371/journal.pcbi.1003731
34. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. (2017) 45:W98–W102. doi: 10.1093/nar/gkx247
35. Gyorffy B, Surowiak P, Budczies J, Lanczky A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS ONE*. (2013) 8:e82241. doi: 10.1371/journal.pone.0082241
36. Xie B, Ding Q, Han H, Wu D. miRCancer: a microRNA-cancer association database constructed by text mining on literature. *Bioinformatics*. (2013) 29:638–44. doi: 10.1093/bioinformatics/btt014
37. Wu G, Ji H. ChIPXpress: using publicly available gene expression data to improve ChIP-seq and ChIP-chip target gene ranking. *BMC Bioinformatics*. (2013) 14:188. doi: 10.1186/1471-2105-14-188
38. Snijders AM, Lee SY, Hang B, Hao W, Bissell MJ, Mao JH. FAM83 family oncogenes are broadly involved in human cancers: an integrative multi-omics approach. *Mol Oncol*. (2017) 11:167–79. doi: 10.1002/1878-0261.12016
39. Leiserson MD, Vandin F, Wu HT, Dobson JR, Eldridge JV, Thomas JL, et al. Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat Genet*. (2015) 47:106–14. doi: 10.1038/ng.3168
40. Ma X, Liu Y, Liu Y, Alexandrov LB, Edmonson MN, Gawad C, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. *Nature*. (2018) 555:371–6. doi: 10.1038/nature25795
41. Ni M, Liu X, Wu J, Zhang D, Tian J, Wang T, et al. Identification of candidate biomarkers correlated with the pathogenesis and prognosis of non-small cell lung cancer via integrated bioinformatics analysis. *Front Genet*. (2018) 9:469. doi: 10.3389/fgene.2018.00469
42. Zhao H, Chen Q, Alam A, Cui J, Suen KC, Soo AP, et al. The role of osteopontin in the progression of solid organ tumour. *Cell Death Dis*. (2018) 9:356. doi: 10.1038/s41419-018-0391-6
43. Cho WY, Hong SH, Singh B, Islam MA, Lee S, Lee AY, et al. Suppression of tumor growth in lung cancer xenograft model mice by poly(sorbitol-co-PEI)-mediated delivery of osteopontin siRNA. *Eur J Pharm Biopharm*. (2015) 94:450–62. doi: 10.1016/j.ejpb.2015.06.017
44. Wang D, Li T, Ye G, Shen Z, Hu Y, Mou T, et al. Overexpression of the receptor for advanced glycation endproducts (RAGE) is associated with poor prognosis in gastric cancer. *PLoS ONE*. (2015) 10:e0122697. doi: 10.1371/journal.pone.0122697
45. Kostova N, Zlateva S, Ugrinova I, Pasheva E. The expression of HMGB1 protein and its receptor RAGE in human malignant tumors. *Mol Cell Biochem*. (2010) 337:251–8. doi: 10.1007/s11010-009-0305-0
46. Wu S, Mao L, Li Y, Yin Y, Yuan W, Chen Y, et al. RAGE may act as a tumour suppressor to regulate lung cancer development. *Gene*. (2018) 651:86–93. doi: 10.1016/j.gene.2018.02.009
47. Meijer B, Geary RB, Day AS. The role of S100A12 as a systemic marker of inflammation. *Int J Inflamm*. (2012) 2012:907078. doi: 10.1155/2012/907078
48. Zhang W, Fan J, Chen Q, Lei C, Qiao B, Liu Q. SPP1 and AGER as potential prognostic biomarkers for lung adenocarcinoma. *Oncol Lett*. (2018) 15:7028–36. doi: 10.3892/ol.2018.8235
49. Langfelder P, Mischel PS, Horvath S. When is hub gene selection better than standard meta-analysis? *PLoS ONE*. (2013) 8:e61505. doi: 10.1371/journal.pone.0061505
50. Barabasi AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet*. (2004) 5:101–13. doi: 10.1038/nrg1272
51. Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Felciano RM, et al. Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc Natl Acad Sci USA*. (2006) 103:17402–7. doi: 10.1073/pnas.0608396103
52. Wang Z, Wan L, Zhong J, Inuzuka H, Liu P, Sarkar FH, et al. Cdc20: a potential novel therapeutic target for cancer treatment. *Curr Pharm Des*. (2013) 19:3210–4. doi: 10.2174/1381612811319180005
53. Nicolau-Neto P, Palumbo A, De Martino M, Esposito F, de Almeida Simao T, Fusco A, et al. UBE2C is a transcriptional target of the cell cycle regulator FOXM1. *Genes (Basel)*. (2018) 9:188. doi: 10.3390/genes9040188
54. Fang Y, Yu H, Liang X, Xu J, Cai X. Chk1-induced CCNB1 overexpression promotes cell proliferation and tumor growth in human colorectal cancer. *Cancer Biol Ther*. (2014) 15:1268–79. doi: 10.4161/cbt.29691
55. Guo C, Wu G, Chin JL, Bauman G, Moussa M, Wang F, et al. Bub1 up-regulation and hyperphosphorylation promote malignant transformation in SV40 tag-induced transgenic mouse models. *Mol Cancer Res*. (2006) 4:957–69. doi: 10.1158/1541-7786.MCR-06-0168
56. Ullmannova V, Popescu NC. Expression profile of the tumor suppressor genes DLC-1 and DLC-2 in solid tumors. *Int J Oncol*. (2006) 29:1127–32. doi: 10.3892/ijo.29.5.1127
57. Yuan BZ, Jefferson AM, Millecchia L, Popescu NC, Reynolds SH. Morphological changes and nuclear translocation of DLC1 tumor suppressor protein precede apoptosis in human non-small cell lung carcinoma cells. *Exp Cell Res*. (2007) 313:3868–80. doi: 10.1016/j.yexcr.2007.08.009
58. Goodison S, Yuan J, Sloan D, Kim R, Li C, Popescu NC, et al. The RhoGAP protein DLC-1 functions as a metastasis suppressor in breast cancer cells. *Cancer Res*. (2005) 65:6042–53. doi: 10.1158/0008-5472.CAN-04-3043
59. Lu J, Xu W, Qian J, Wang S, Zhang B, Zhang L, et al. Transcriptome profiling analysis reveals that CXCL2 is involved in anlotinib resistance in human lung cancer cells. *BMC Med Genomics*. (2019) 12(Suppl. 2):38. doi: 10.1186/s12920-019-0482-y
60. Wei S, Dunn TA, Isaacs WB, De Marzo AM, Luo J. GOLPH2 and MYO6: putative prostate cancer markers localized to the Golgi apparatus. *Prostate*. (2008) 68:1387–95. doi: 10.1002/pros.20806
61. Zhang F, Gu Y, Li X, Wang W, He J, Peng T. Up-regulated Golgi phosphoprotein 2 (GOLPH2) expression in lung adenocarcinoma tissue. *Clin Biochem*. (2010) 43:983–91. doi: 10.1016/j.clinbiochem.2010.05.010
62. Litovchick L, Sadasivam S, Florens L, Zhu X, Swanson SK, Velmurugan S, et al. Evolutionarily conserved multisubunit RBL2/p130 and E2F4 protein complex represses human cell cycle-dependent genes in quiescence. *Mol Cell*. (2007) 26:539–51. doi: 10.1016/j.molcel.2007.04.015
63. Fischer M, Grossmann P, Padi M, DeCaprio JA. Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F target gene analyses identifies cell cycle gene regulatory networks. *Nucleic Acids Res*. (2016) 44:6070–86. doi: 10.1093/nar/gkw523
64. Iness AN, Felthousen J, Ananthapadmanabhan V, Sesay F, Saini S, Guiley KZ, et al. The cell cycle regulatory DREAM complex is disrupted by high expression of oncogenic B-Myb. *Oncogene*. (2019) 38:1080–92. doi: 10.1038/s41388-018-0490-y
65. Zhang J, Zhang J, Cui X, Yang Y, Li M, Qu J, et al. FoxM1: a novel tumor biomarker of lung cancer. *Int J Clin Exp Med*. (2015) 8:3136–40.
66. Wang IC, Meliton L, Ren X, Zhang Y, Balli D, Snyder J, et al. Deletion of Forkhead Box M1 transcription factor from respiratory epithelial cells inhibits pulmonary tumorigenesis. *PLoS ONE*. (2009) 4:e6609. doi: 10.1371/journal.pone.0006609
67. McClendon AK, Osheroff N. DNA topoisomerase II, genotoxicity, and cancer. *Mutat Res*. (2007) 623:83–97. doi: 10.1016/j.mrfmmm.2007.06.009
68. Barger CJ, Branick C, Chee L, Karpf AR. Pan-cancer analyses reveal genomic features of FOXM1 overexpression in cancer. *Cancers (Basel)*. (2019) 11:E251. doi: 10.3390/cancers11020251



69. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med.* (2015) 21:938–45. doi: 10.1038/nm.3909
70. Fan X, Wang Y, Jiang T, Cai W, Jin Y, Niu Y, et al. B-Myb mediates proliferation and migration of non-small-cell lung cancer via suppressing IGFBP3. *Int J Mol Sci.* (2018) 19:1479. doi: 10.3390/ijms19051479
71. Souza RF, Yin J, Smolinski KN, Zou TT, Wang S, Shi YQ, et al. Frequent mutation of the E2F-4 cell cycle gene in primary human gastrointestinal tumors. *Cancer Res.* (1997) 57:2350–3.
72. Schwemmler S, Pfeifer GP. Genomic structure and mutation screening of the E2F4 gene in human tumors. *Int J Cancer.* (2000) 86:672–7.
73. Shimo A, Tanikawa C, Nishidate T, Lin ML, Matsuda K, Park JH, et al. Involvement of kinesin family member 2C/mitotic centromere-associated kinesin overexpression in mammary carcinogenesis. *Cancer Sci.* (2008) 99:62–70. doi: 10.1111/j.1349-7006.2007.00635.x
74. Narayan G, Bourdon V, Chaganti S, Arias-Pulido H, Nandula SV, Rao PH, et al. Gene dosage alterations revealed by cDNA microarray analysis in cervical cancer: identification of candidate amplified and overexpressed genes. *Genes Chromosomes Cancer.* (2007) 46:373–84. doi: 10.1002/gcc.20418
75. Castillo A, Morse HC, 3rd, Godfrey VL, Naeem R, Justice MJ. Overexpression of Eg5 causes genomic instability and tumor formation in mice. *Cancer Res.* (2007) 67:10138–47. doi: 10.1158/0008-5472.CAN-07-0326
76. Mazumdar M, Sundareshan S, Misteli T. Human chromokinesin KIF4A functions in chromosome condensation and segregation. *J Cell Biol.* (2004) 166:613–20. doi: 10.1083/jcb.200401142
77. Yu Y, Feng YM. The role of kinesin family proteins in tumorigenesis and progression: potential biomarkers and molecular targets for cancer therapy. *Cancer.* (2010) 116:5150–60. doi: 10.1002/ncr.25461
78. Xu Z, Zhang Q, Luh F, Jin B, Liu X. Overexpression of the ASPM gene is associated with aggressiveness and poor outcome in bladder cancer. *Oncol Lett.* (2019) 17:1865–76. doi: 10.3892/ol.2018.9762
79. Zhong X, Guan X, Dong Q, Yang S, Liu W, Zhang L. Examining Nek2 as a better proliferation marker in non-small cell lung cancer prognosis. *Tumour Biol.* (2014) 35:7155–62. doi: 10.1007/s13277-014-1935-8
80. Fang Y, Zhang X. Targeting NEK2 as a promising therapeutic approach for cancer treatment. *Cell Cycle.* (2016) 15:895–907. doi: 10.1080/15384101.2016.1152430
81. Hu P, Shanguan J, Zhang L. Downregulation of NUF2 inhibits tumor growth and induces apoptosis by regulating lncRNA AF339813. *Int J Clin Exp Pathol.* (2015) 8:2638–48.
82. O'Brien SL, Fagan A, Fox EJ, Millikan RC, Culhane AC, Brennan DJ, et al. CENP-F expression is associated with poor prognosis and chromosomal instability in patients with primary breast cancer. *Int J Cancer.* (2007) 120:1434–43. doi: 10.1002/ijc.22413
83. Kim HE, Kim DG, Lee KJ, Son JG, Song MY, Park YM, et al. Frequent amplification of CENPF, GMNN and CDK13 genes in hepatocellular carcinomas. *PLoS ONE.* (2012) 7:e43223. doi: 10.1371/journal.pone.0043223
84. Ahmed F, Raghava GP. Designing of highly effective complementary and mismatch siRNAs for silencing a gene. *PLoS ONE.* (2011) 6:e23443. doi: 10.1371/journal.pone.0023443
85. Li J, Wang Y, Luo J, Fu Z, Ying J, Yu Y, et al. miR-134 inhibits epithelial to mesenchymal transition by targeting FOXM1 in non-small cell lung cancer cells. *FEBS Lett.* (2012) 586:3761–5. doi: 10.1016/j.febslet.2012.09.016
86. Ke Y, Zhao W, Xiong J, Cao R. miR-149 inhibits non-small-cell lung cancer cells EMT by targeting FOXM1. *Biochem Res Int.* (2013) 2013:506731. doi: 10.1155/2013/506731
87. Zhu X, Li D, Yu F, Jia C, Xie J, Ma Y, et al. miR-194 inhibits the proliferation, invasion, migration, and enhances the chemosensitivity of non-small cell lung cancer cells by targeting forkhead box A1 protein. *Oncotarget.* (2016) 7:13139–52. doi: 10.18632/oncotarget.7545
88. Wang P, Lv HY, Zhou DM, Zhang EN. miR-204 suppresses non-small-cell lung carcinoma (NSCLC) invasion and migration by targeting JAK2. *Genet Mol Res.* (2016) 2:gmr.15026415. doi: 10.4238/gmr.15026415
89. Wang Y, Chen J, Lin Z, Cao J, Huang H, Jiang Y, et al. Role of deregulated microRNAs in non-small cell lung cancer progression using fresh-frozen and formalin-fixed, paraffin-embedded samples. *Oncol Lett.* (2016) 11:801–8. doi: 10.3892/ol.2015.3976
90. Zhang G, Jiang G, Wang C, Zhong K, Zhang J, Xue Q, et al. Decreased expression of microRNA-320a promotes proliferation and invasion of non-small cell lung cancer cells by increasing VDACL1 expression. *Oncotarget.* (2016) 7:49470–80. doi: 10.18632/oncotarget.9943
91. Chen W, Wang J, Liu S, Wang S, Cheng Y, Zhou W, et al. MicroRNA-361-3p suppresses tumor cell proliferation and metastasis by directly targeting SH2B1 in NSCLC. *J Exp Clin Cancer Res.* (2016) 35:76. doi: 10.1186/s13046-016-0357-4
92. Chen MJ, Wu DW, Wang GC, Wang YC, Chen CY, Lee H. MicroRNA-630 may confer favorable cisplatin-based chemotherapy and clinical outcomes in non-small cell lung cancer by targeting Bcl-2. *Oncotarget.* (2018) 9:13758–67. doi: 10.18632/oncotarget.24474
93. Geng GJ, Yang YT, Jiang J, Yu XY, Fa XE. MicroRNA-30a suppresses non-small-cell lung cancer by targeting Myb-related protein B. *Exp Ther Med.* (2018) 15:1633–9. doi: 10.3892/etm.2017.5559
94. Huang R, Gao L. Identification of potential diagnostic and prognostic biomarkers in non-small cell lung cancer based on microarray data. *Oncol Lett.* (2018) 15:6436–42. doi: 10.3892/ol.2018.8153
95. Piao J, Sun J, Yang Y, Jin T, Chen L, Lin Z. Target gene screening and evaluation of prognostic values in non-small cell lung cancers by bioinformatics analysis. *Gene.* (2018) 647:306–11. doi: 10.1016/j.gene.2018.01.003
96. Ahmed F, Benedito VA, Zhao PX. Mining functional elements in messenger RNAs: overview, challenges, and perspectives. *Front Plant Sci.* (2011) 2:84. doi: 10.3389/fpls.2011.00084
97. Ahmed F, Kumar M, Raghava GP. Prediction of polyadenylation signals in human DNA sequences using nucleotide frequencies. *In Silico Biol.* (2009) 9:135–48.
98. Landi D, Gemignani F, Landi S. Role of variations within microRNA-binding sites in cancer. *Mutagenesis.* (2012) 27:205–10. doi: 10.1093/mutage/ger055

**Conflict of Interest:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Ahmed. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.