

# Weighted gene coexpression network analysis identifies hub genes related to KRAS mutant lung adenocarcinoma

Dongjun Dai, PhD<sup>a</sup>, Rongkai Shi, Master degree<sup>a</sup>, Shuting Han, PhD<sup>a</sup>, Hongchuan Jin, PhD<sup>b</sup>, Xian Wang, PhD<sup>a,\*</sup>

## Abstract

The aim of current study was to use Weighted Gene Coexpression Network Analysis (WGCNA) to identify hub genes related to the incidence and prognosis of KRAS mutant (MT) lung adenocarcinoma (LUAD).

We involved 184 stage IIB to IV LUAD samples and 59 normal lung tissue samples from The Cancer Genome Atlas (TCGA) database. The R package “limma” was used to identify differentially expressed genes (DEGs). WGCNA and survival analyses were performed by R packages “WGCNA” and “survival,” respectively. The functional analyses were performed by R package “clusterProfiler” and GSEA software. Network construction and MCODE analysis were performed by Cytoscape\_v3.6.1.

Totally 2590 KRAS MT specific DEGs were found between LUAD and normal lung tissues, and 10 WGCNA modules were identified. Functional analysis of the key module showed the ribosome biogenesis related terms were enriched. We observed the expression of 8 genes were positively correlated to the worse survival of KRAS MT LUAD patients, the 7 of them were validated by Kaplan–Meier plotter database (kmpplot.com/) (thymosin Beta 10 [TMSB10], ribosomal Protein S16 [RPS16], mitochondrial ribosomal protein L27 [MRPL27], cytochrome c oxidase subunit 6A1 [COX6A1], HCLS1-associated protein X-1 [HAX1], ribosomal protein L38 [RPL38], and ATP Synthase Membrane Subunit DAPIT [ATP5MD]). The GSEA analysis found mTOR and STK33 pathways were upregulated in KRAS MT LUAD ( $P < .05$ , false discovery rate [FDR]  $< 0.25$ ).

In summary, our study firstly used WGCNA to identify hub genes in the development of KRAS MT LUAD. The identified prognostic factors would be potential biomarkers in clinical use. Further molecular studies are required to confirm the mechanism of those genes in KRAS MT LUAD.

**Abbreviations:** 4-HPR = N-(4-hydroxyphenyl) retinamide, COX = Cytochrome oxidase, DEGs = differentially expressed genes, F-ATPase = Fo subunit of the mitochondrial H<sup>+</sup>-ATP synthase, FDR = false discovery rate, FC = fold change, FPKM = Fragments per Kilobase Million, GAPs = GTPase activating proteins, GO = Gene Ontology, GS = gene significance, GSEA = gene set enrichment analysis, KEGG = Kyoto Encyclopedia of Genes and Genomes (KEGG), KM = Kaplan–Meier, KMPLOT = Kaplan–Meier plotter database, KRAS MT = KRAS mutant, LUAD = lung adenocarcinoma, MCODE = molecular complex detection, ME = module eigengene, MM = Module Membership, MS = module significance, NSCLC = non-small-cell lung cancer, OS = overall survival, PFS = progression-free survival, RNA-seq = RNA-sequencing, TCGA = The Cancer Genome Atlas, TOM = topological overlap matrix, V-ATPase = Vacuolar proton pump, WGCNA = Weighted Gene Coexpression Network Analysis.

**Keywords:** hub gene, KRAS mutant lung adenocarcinoma, prognosis, The Cancer Genome Atlas, Weighted Gene Coexpression Network Analysis

Editor: Balaji Thas Moorthy.

This work was supported by National Natural Science Foundation Of China (81502386; 81772944) and Zhejiang Provincial Program for High-level Innovative Healthcare talents.

The authors have no conflicts of interest to disclose.

Supplemental Digital Content is available for this article.

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

<sup>a</sup> Department of Medical Oncology, <sup>b</sup> Laboratory of Cancer Biology, Key Lab of Biotherapy, Sir Run Run Shaw Hospital, Medical School of Zhejiang University, Hangzhou, China.

\* Correspondence: Xian Wang, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China (e-mail: wangx118@zju.edu.cn).

Copyright © 2020 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Dai D, Shi R, Han S, Jin H, Wang X. Weighted gene co-expression network analysis identifies hub genes related to KRAS mutant lung adenocarcinoma. *Medicine* 2020;99:32(e21478).

Received: 28 February 2020 / Received in final form: 20 June 2020 / Accepted: 26 June 2020

<http://dx.doi.org/10.1097/MD.00000000000021478>

## 1. Introduction

Lung cancer is the leading cause of cancer related deaths worldwide.<sup>[1]</sup> Non-small-cell lung cancer (NSCLC) accounts for 80% of all lung cancer and is often diagnosed at an advanced stage, which lead to a poor prognosis.<sup>[2]</sup> The main type of NSCLC is lung adenocarcinoma (LUAD).<sup>[3]</sup>

The RAS gene family includes KRAS, NRAS, and HRAS, which encodes for plasma membrane-localized transduction proteins with intrinsic GTPase activity. Normally, the intrinsic GTPase activity of RAS proteins is accelerated by GTPase activating proteins (GAPs), which attenuate GTPase signaling by accelerating the conversion of KRAS-GTP bound status to KRAS-GDP bound status. However, the mutation of RAS gene leads to deficiency of GTPase degradation which results in a constitutively KRAS-GTP bound state and in turn activates the downstream pathways to promote cell proliferation.<sup>[4]</sup> The mutation of RAS gene was found in approximately 30% of human cancers.<sup>[5]</sup> RAS mutation was observed to promote many types of cancers.<sup>[5]</sup> RAS mutation is implicated with the cancer development and the resistance to cancer therapies, such as EGFR inhibitors treatment and chemotherapy.<sup>[6]</sup> KRAS was identified as an outstanding predictive biomarker for LUAD, which generally associated with worse progression-free survival (PFS) and overall survival (OS).<sup>[7]</sup>

RAS protein is famous for its “undruggable.” No effective direct RAS inhibitor has been approved.<sup>[8]</sup> In KRAS mutant (KRAS MT) LUAD, many indirect strategies were developed to target the downstream pathways of KRAS such as RAF/MAPK/ERK pathway, PI3K/AKT/mTOR pathway, and RHO-FAK pathway.<sup>[9]</sup> The KRAS downstream targets remain worth to be discovered. Weighted Gene Coexpression Network Analysis (WGCNA) is a valuable modular biological analysis methodology, which could identify and screen co-expressed gene modules and key biomarkers.<sup>[10]</sup> The WGCNA is a suitable tool for identifying new hub genes and pathways in the downstream of KRAS pathway. To our best knowledge, this approach has not been applied in KRAS MT LUAD. The aim of our study was to identify novel coexpression gene network modules associated with KRAS MT LUAD by WGCNA, and in turn to find potential signal pathways and genes that involved in the pathogenesis and prognosis of KRAS MT LUAD.

## 2. Methods

### 2.1. Data collection

We downloaded our data from The Cancer Genome Atlas (TCGA) database.<sup>[11]</sup> As the 5-year survival rate dropped sharply in higher stage lung cancer patients,<sup>[12]</sup> we included 184 stage IIB to IV LUAD samples and 59 normal lung tissue samples with clinical information and RNA-sequencing (RNA-seq) data. The read counts and Fragments per Kilobase Million (FPKM) data were both collected. Ethical approval was not available for the study as our data were revealed from public database.

### 2.2. Differentially expressed genes (DEGs) screening

The read counts data were used to screen the DEGs between LUAD samples and normal tissues by using R package “limma.” The genes with false discovery rate (FDR) < 0.05 and  $|\log_2$  fold change (FC)|  $\geq 0.585$  were selected for further analysis. As shown in Fig. 1, we first screened DEGs between KRAS wild type LUAD or KRAS MT LUAD samples and normal tissues, respectively.

Next, we selected KRAS MT specific DEGs by removing the intersect DEGs of these 2 groups in the KRAS MT group. The FPKM data of these genes were further analyzed by WGCNA.

### 2.3. Weighted gene coexpression networks construction and module selection

WGCNA was performed by R package “WGCNA.”<sup>[10]</sup> The KRAS MT LUAD samples (N=56) and the normal tissue (N=59) samples were included in the WGCNA analysis. Pearson’s correlation matrices were constructed for pair-wise genes. A weighted adjacency matrix of genes using a power function  $am_n = |cm_n|^\beta$  ( $cm_n$  = Pearson correlation between gene m and gene n;  $am_n$  = adjacency between gene m and gene n) was built. The  $\beta$  was a softthresholding-power parameter which could strengthen strong correlations and penalize the weak correlation. The best  $\beta$  here was 7, with a scale free  $R^2 = 0.82$  (Supplemental Digital Content [Figure S1, <http://links.lww.com/MD/E640>]). Next, the adjacency was transformed into topological overlap matrix (TOM) to measure the network connectivity of single gene, which was defined as the sum of its adjacency with all other genes for network generation. Then, an average linkage hierarchical clustering was built to find modules with genes of similar expression profile by using the TOM-based dissimilarity measure with a minimum group size of 40 genes for dendrogram.

### 2.4. Identification of significant modules related to KRAS MT LUAD

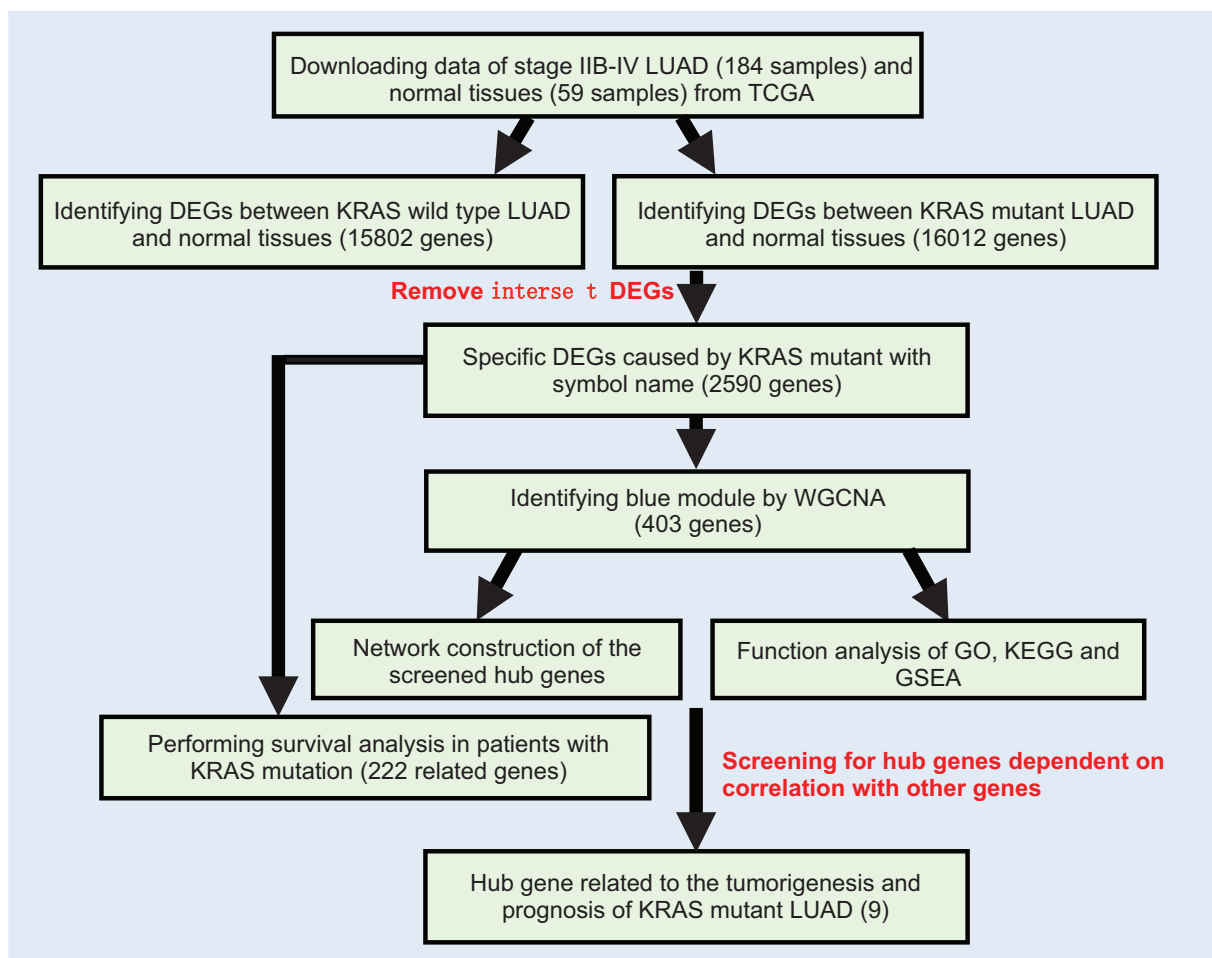
The module eigengene (ME) was calculated for representing and summarizing each module. The ME was defined as the first principal component of a given module, which could be representative of the gene expression profile in a module. The correlation between KRAS MT LUAD and MEs was calculated to identify the key module related to KRAS MT LUAD. Gene significance (GS) represents the absolute value of the correlation between a specific gene and the KRAS MT LUAD. Module significance (MS) was defined as the average GS for all genes in a module. The module with the highest MS was considered as the key module related with KRAS MT LUAD.

### 2.5. Analysis of hub genes in the most significant module

Hub genes are the highly connected nodes with more interaction than other genes in a module. For intramodular analysis, the GS and module membership (MM, which is also known as eigengene-based connectivity kME) were calculated. Genes with higher MM and GS in the key module would be defined as hub genes. Here we selected the hub gene by external traits based  $|GS| > 0.2$  and  $|MM| > 0.7$  with a threshold of  $P$ -value < .05. MCODE (molecular complex detection) analysis was performed to select the sub-module, the parameters were set as the follows: degree cut-of=2, node score cut-of=0.2, k-core=2, and max. depth=100. The hub genes interaction network and MCODE analysis were performed by Cytoscape\_v3.6.1.

### 2.6. Functional enrichment analysis

The functional analysis was performed by Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO) analyses, and gene set enrichment analysis (GSEA). The KEGG and GO analyses were conducted by R package “clusterProfiler.”<sup>[13]</sup> The



**Figure 1.** The workflow of current study. Legends: Gene dendrogram was obtained by average linkage hierarchical clustering. The color row underneath the dendrogram indicates module membership.

GSEA was performed by the GSEA software from Broad Institute (<http://software.broadinstitute.org/gsea/downloads.jsp>). The  $P < .05$  and  $FDR < 0.25$  were set as the cut-off criterion.

### 2.7. Survival analysis of DEGs in KRAS mutation patients

We performed survival analysis of each KRAS MT specific DEGs by using R package “survival.” The patients were divided into 2 groups based on whether its expression was over the median value of all samples. The univariate Cox proportional hazards regression model was used to find survival related genes. The screened survival related genes overlapped with the hub genes of MCODE analysis were considered as hub survival genes. The Kaplan–Meier (KM) method was applied to draw survival curve. Log-rank  $P$ -value was calculated for KM plot. The validation of the hub survival genes was performed using the Kaplan–Meier plotter database (KMPLLOT) (<http://kmplot.com/analysis/>).

## 3. Results

### 3.1. Screening the mRNA modules

As shown in Fig. 1, we identified 15,802 DEGs between KRAS wild type LUAD and normal lung tissues and 16,012 DEGs

between KRAS MT LUAD and normal lung tissues, respectively. After removing the intersect DEGs, we involved 2590 genes into the WGCNA (1606 upregulated genes and 984 downregulated genes in KRAS MT LUAD tissue, Supplemental Digital Content [Table S1, <http://links.lww.com/MD/E645>]). After WGCNA analysis, we revealed 10 modules (“MEagenta,” “MEbrown,” “MEpink,” “MEblack,” “MEblue,” “MEturquoise,” “MEyellow,” “MEgreen,” “MERed,” “MEgrey,” Table 1, Supplemental Digital Content [Figure S2, <http://links.lww.com/MD/E641>]). The heatmap of the correlation between MEs and KRAS MT LUAD showed that module blue had the strongest correlation with KRAS MT LUAD (Fig. 2A). We found module blue also had the highest MS score (Fig. 2B). A scatterplot found GS and MM exhibited a very significant correlation between module blue and KRAS MT LUAD (Fig. 2C, Table 1). Furthermore, hierarchical clustering dendrogram showed the KRAS MT LUAD was highly related to module blue, as evidenced by their low merging height (Fig. 2D). Those results implied that module blue was the key module in the KRAS MT LUAD.

### 3.2. Functional analysis of blue module

Functional analyses were performed for genes in module blue to clarify its biological meaning. GO analysis was applied from 3

**Table 1**  
The modules identified from WGCNA for KRAS MT LUAD.

Modules	Total gene count	Correlation with tumor stage	
		R <sup>2</sup>	P-value
Black	67	-0.23	.061
Blue	403	0.68	5.30E-56
Brown	236	0.2	.002
Green	95	0.09	.39
Magenta	51	-0.42	.0022
Pink	61	-0.4	.0014
Red	81	-0.26	.019
Turquoise	455	0.29	2.90E-10
Yellow	130	0.18	.04
Grey	1011	NA	NA

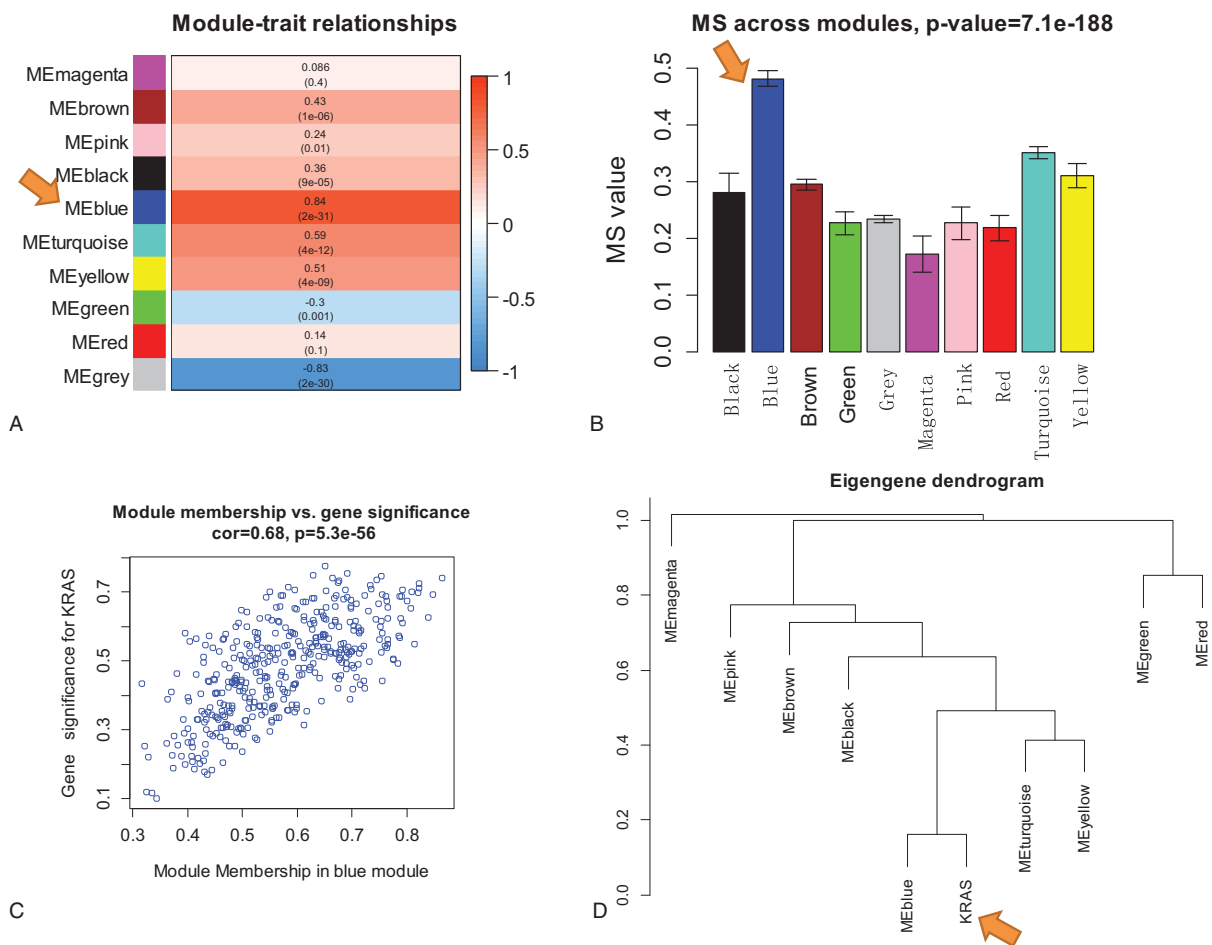
To be noted, the grey module belongs to WGCNA special category that has no association with the clinical trait analyzed. NA=not available.

aspects: biological process (BP), cellular component (CC), and molecular function (MF). Both the GO and KEGG analyses showed the proportion of genes increased significantly in ribosome and translation terms (Fig. 3, Supplemental Digital

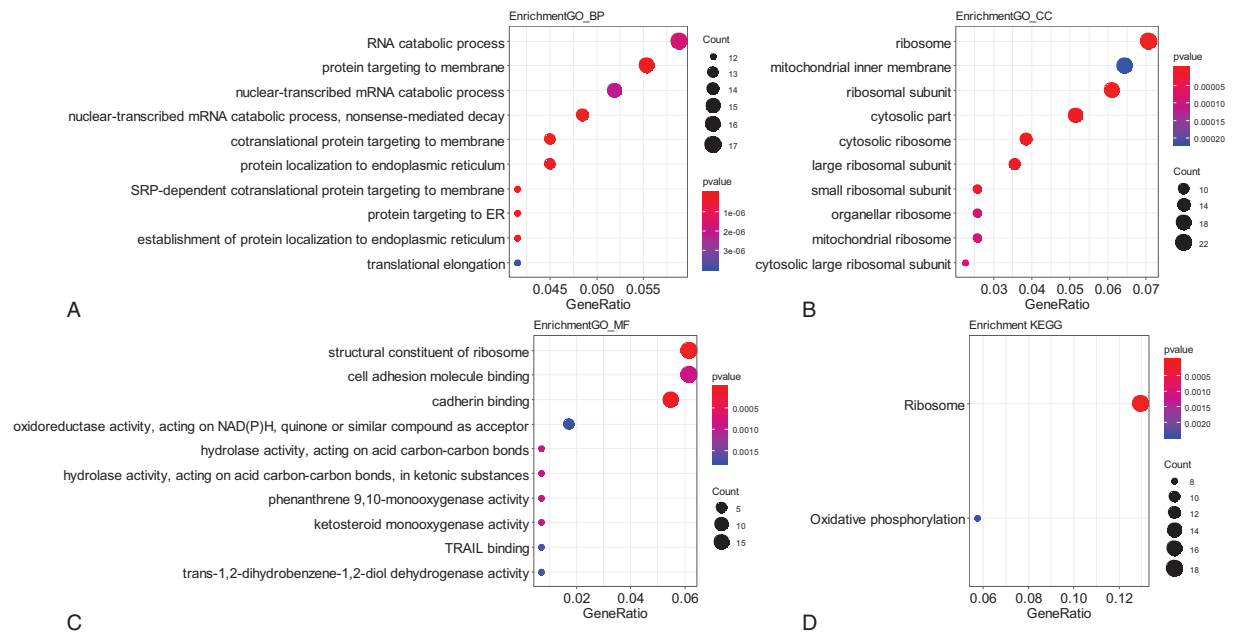
Content [Table S2, <http://links.lww.com/MD/E646>, S3, <http://links.lww.com/MD/E647>, S4, <http://links.lww.com/MD/E648>, and S5, <http://links.lww.com/MD/E649>]). Further GSEA analysis of GO terms showed the ribosome related gene sets were enriched in KRAS MT LUAD tissues (Supplemental Digital Content [Table S6, <http://links.lww.com/MD/E650>]). Besides, GSEA analysis of oncogenic signatures showed mTOR<sup>[14]</sup> and STK33<sup>[15]</sup> pathways were enriched in KRAS MT LUAD tissues (Supplemental Digital Content [Table S7, <http://links.lww.com/MD/E651>]).

### 3.3. Detection of hub gene and their functional annotations

Hub genes selection was performed by the criteria mentioned in method section. There were 64 significant genes with |GS| > 0.2 and |MM| > 0.7. After MCODE analysis, a network of 57 genes was constructed (Supplemental Digital Content [Figure S3, <http://links.lww.com/MD/E642>], Table 2). Furthermore, as KRAS was identified as a poor prognostic predictor in lung cancer,<sup>[16]</sup> we performed survival analysis between each gene of the KRAS MT specific DEGs in KRAS MT LUAD patients. There were 222



**Figure 2.** The identification of key module related to KRAS MT LUAD. Legends: A, Heatmap of the correlation between module eigengenes and KRAS MT LUAD; B, MS values of each module associated with KRAS MT LUAD; C, A scatterplot of gene significance for KRAS MT LUAD versus module membership in module blue. GS and MM exhibit a very significant correlation, implying that hub genes of the module blue also tend to be highly correlated with KRAS MT LUAD incidence; D, Hierarchical clustering dendrogram of module eigengenes. The KRAS MT LUAD and module blue eigengenes are highly related, as evidenced by their low merging height. KRAS MT =KRAS mutant, LUAD=lung adenocarcinoma.



**Figure 3.** Functional analysis of module blue. Legends: A–C, Gene ontology enrichment analysis for biological process (BP). A, cellular component (CC) B, and molecular function (MF) C, of genes in module blue; D, KEGG pathway enrichment analysis of genes in module blue; The size of the bubble indicated the counts of genes, and the color indicated the *P*-value. KEGG=Kyoto Encyclopedia of Genes and Genomes.

survival related genes screened (70 genes positively correlated and 152 genes negatively correlated with worse KRAS MT LUAD prognosis, respectively, Supplemental Digital Content [Table S8, <http://links.lww.com/MD/E652>]). Among them, we found 8 genes, including thymosin Beta 10 (TMSB10), ribosomal protein L26 like 1 (RPL26L1), ribosomal Protein S16 (RPS16), mitochondrial ribosomal protein L27 (MRPL27), cytochrome c oxidase subunit 6A1 (COX6A1), HCLS1-associated protein X-1 (HAX1), ribosomal protein L38 (RPL38), and ATP Synthase Membrane Subunit DAPIT (ATP5MD), were overlapped with the hub genes in MCODE module and were positively correlated with worse prognosis of KRAS MT LUAD (Fig. 4). The expression of these 8 genes were observed to be higher in KRAS MT LUAD tissues (Supplemental Digital Content [Figure S4, <http://links.lww.com/MD/E643>]). Moreover, the validation of those prognostic genes by KMPLOT showed consistent results except for RPL26L1 (Supplemental Digital Content [Figure S5, <http://links.lww.com/MD/E644>]).

#### 4. Discussion

The WGCNA is a useful tool to find modules of highly correlated genes. The intramodular hub genes of the key module could be used for the diagnostic and prognostic prediction of disease such as cancer.<sup>[17]</sup> We performed WGCNA for stage IIB to IV KRAS MT LUAD and normal lung samples using specific DEGs caused by KRAS mutation. The calculation of correlation between coexpression module and KRAS MT LUAD identified module blue as the key module of tumorigenesis of KRAS MT LUAD. Functional analyses found ribosome related terms, mTOR and STK33 pathways were enriched in KRAS MT LUAD. We further determined hub genes in module blue which were related to KRAS MT LUAD. Moreover, there were 8 hub genes correlated with the prognosis of KRAS MT LUAD patients.

The hyperactivation of ribosome biogenesis has a critical role in the initiation and progression of cancers.<sup>[18]</sup> The alteration of ribosome biogenesis influences the translation of cancer genome. Major cancer related signaling pathways, such as MYC, RAS–MAPK, PI3K–AKT–mTOR, and WNT– $\beta$ -catenin, were found to be involved in the aberrant activation of mRNA translation at the initiation and elongation steps.<sup>[18]</sup> Drugs had been discovered to inhibit ribosome biogenesis to offer a viable therapeutic approach for cancer treatment.<sup>[19]</sup> In our study, ribosome biogenesis related gene sets were found to be upregulated in KRAS MT LUAD tissues. In addition, we also identified ribosome related genes RPS16, RPL26L1, MRPL27, and RPL38 as hub genes and prognostic markers in KRAS MT LUAD. The mRNA expression of those genes were also proved to be upregulated in cancers by previous studies.<sup>[20–23]</sup> Furthermore, PI3K–AKT–mTOR pathway could cause hyperactivation of RNA polymerase I-dependent transcription and ribosome biogenesis, which would lead to sustain proliferation of cancer cells through the upregulation of global protein synthesis.<sup>[24]</sup> Our GSEA analysis found mTOR pathway was enriched in KRAS MT LUAD tissues. The current WGCNA study proved new hint to the relationship between ribosome/translation and the LUAD.

Multiple KRAS MT cancer cell lines were found to be sensitive to the suppression of STK33. The STK33 could suppress the mitochondrial apoptosis in KRAS MT cell by Ribosomal protein S6 kinase beta-1 (S6K1)-induced inactivation of the death agonist BAD. The STK33 was identified as a drug target for the treatment of KRAS MT cancer cell.<sup>[15]</sup> S6K1 is a serine/threonine kinase in the downstream of mTOR pathway. The target substrate of S6K1 is the S6 ribosomal protein.<sup>[25]</sup> The phosphorylation of S6 ribosomal protein induces protein synthesis at the ribosome, which is consistent with our findings that the mTOR pathway and the ribosome related pathway were upregulated in KRAS MT LUAD.

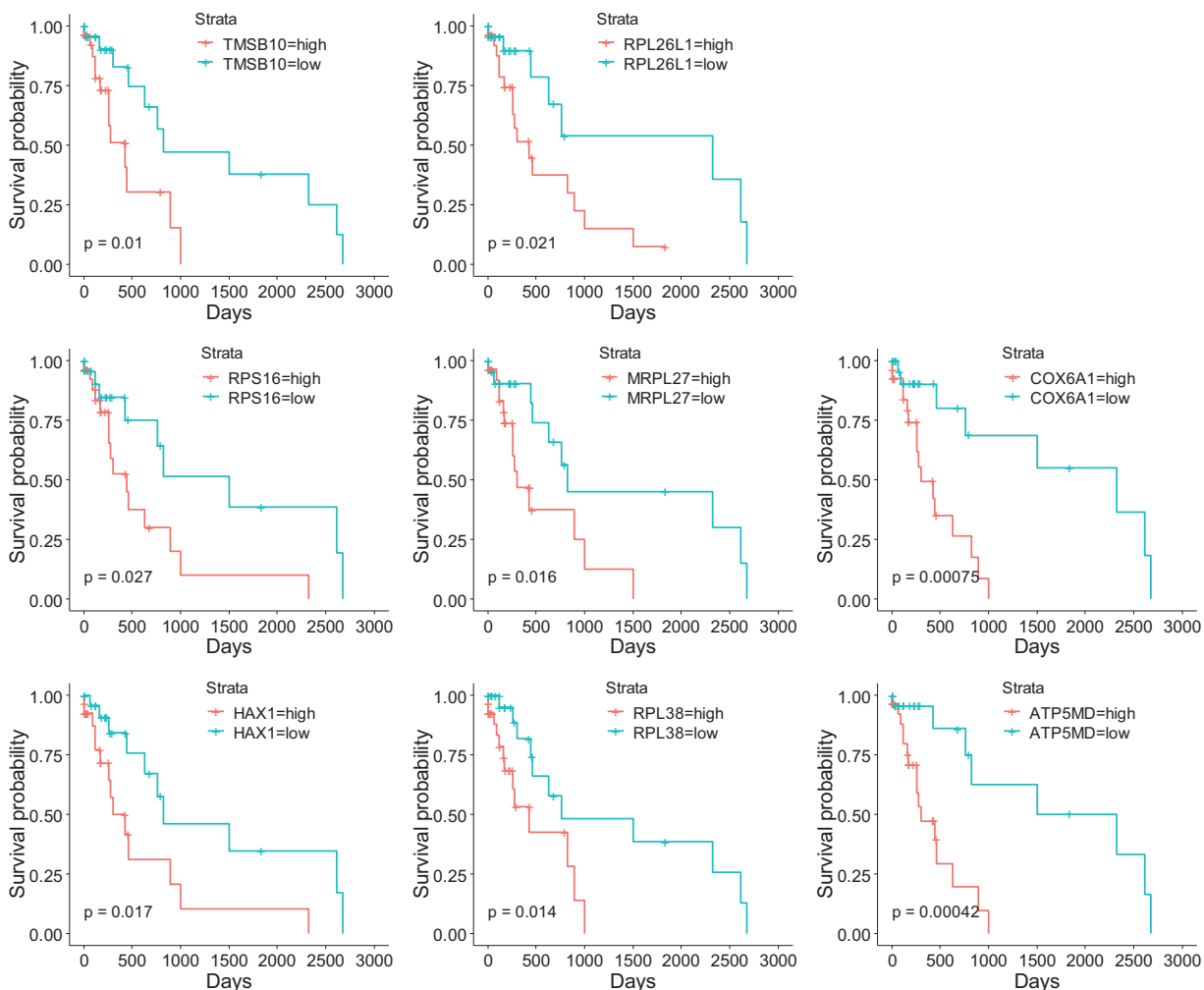
**Table 2****The hub genes selected after MCODE analysis.**

Name	MM P value	MM value	GS value	MCODE_Score
ACBD6	2.56E-29	0.82	0.71	44.71
APLP2	3.92E-19	-0.71	-0.58	43.65
ATP5MD	1.89E-22	0.76	0.44	44.71
ATP5PD	3.89E-22	0.75	0.60	44.71
ATXN1	2.78E-19	-0.72	-0.43	42.72
AUP1	1.84E-35	0.86	0.74	44.71
C1orf43	2.44E-28	0.81	0.55	44.71
CCDC51	5.72E-19	0.71	0.49	42.83
CCT7	9.28E-27	0.80	0.67	44.71
CDK7	7.25E-22	0.75	0.67	44.71
CHCHD1	1.54E-31	0.84	0.63	44.71
COX6A1	2.04E-18	0.70	0.48	42.94
COX6B1	2.79E-23	0.76	0.55	44.71
CYLD	6.30E-23	-0.76	-0.54	44.71
FBXO22	2.05E-19	0.72	0.54	43.25
FBXO7	9.47E-21	-0.73	-0.71	42.83
FCHSD2	1.28E-27	-0.81	-0.65	44.71
FH	3.16E-23	0.76	0.67	40.49
GOLIM4	3.11E-18	-0.70	-0.49	39.95
GUCD1	3.13E-22	-0.75	-0.74	44.05
HAX1	3.23E-24	0.77	0.51	44.71
HNRNPC	8.40E-26	0.79	0.69	44.71
IMPDH2	2.08E-18	0.70	0.59	43.65
KDELRL1	1.24E-21	0.75	0.74	43.80
LAMTOR2	2.08E-29	0.82	0.72	44.71
MAIP1	1.27E-25	0.79	0.62	44.71
MRPL27	1.91E-25	0.79	0.57	44.71
MRPL35	2.67E-23	0.76	0.66	44.71
MRPL44	2.33E-21	0.74	0.62	44.71
MYDGF	2.71E-23	0.76	0.62	44.71
NDUFAF2	4.52E-19	0.71	0.52	44.71
NDUFS2	5.56E-26	0.79	0.71	40.09
ODR4	5.02E-23	0.76	0.69	36.22
ORMDL2	9.77E-20	0.72	0.54	43.92
OST4	7.11E-26	0.79	0.61	44.71
PIGF	3.91E-19	0.71	0.59	43.10
PPP2R5B	2.14E-19	-0.72	-0.59	41.83
RNF146	1.17E-18	-0.71	-0.62	37.81
RPL12	1.77E-20	0.73	0.50	44.29
RPL23A	3.77E-23	0.76	0.54	44.71
RPL26L1	2.31E-23	0.77	0.56	44.71
RPL29	2.76E-18	0.70	0.42	43.39
RPL37A	2.95E-24	0.77	0.50	44.71
RPL38	9.79E-19	0.71	0.45	44.29
RPS16	5.45E-25	0.78	0.49	44.71
RPS19	1.03E-19	0.72	0.44	44.29
RSL1D1	1.88E-25	0.79	0.69	43.76
SLIRP	2.81E-22	0.75	0.53	44.71
SRP9	2.94E-22	0.75	0.61	44.71
SSR2	2.85E-25	0.79	0.57	44.71
TESK2	9.84E-28	-0.81	-0.70	43.03
TMC01	1.06E-32	0.85	0.69	44.71
TMEM183A	6.04E-22	0.75	0.56	44.71
TMSB10	1.90E-18	0.70	0.60	44.58
UFC1	1.81E-22	0.76	0.58	44.71
UQCRQ	8.40E-20	0.72	0.42	44.29
ZNF136	3.02E-29	-0.82	-0.70	44.71

GS= gene significance, MM= module membership.

TMSB10 is a monomeric actin sequestering protein, which is involved in cell motility. The TMSB10 protein was found to be highly expressed in most of the cancer types including the lung cancer.<sup>[26]</sup> TMSB10 could inhibit the apoptosis and promote the

proliferation of lung cancer.<sup>[27]</sup> However, controversial results also exist in the expression of TMSB10 in lung cancer tissue.<sup>[26]</sup> Our results showed TMSB10 was upregulated in KRAS MT LUAD tissues and promoted the tumor progression of KRAS MT



**Figure 4.** The KM plots of hub genes related to survival of KRAS MT LUAD patients. KM = Kaplan–Meier, KRAS MT = KRAS mutant, LUAD = lung adenocarcinoma.

LUAD patients, which provided new evidence for the relationship between TMSB10 and lung cancer.

HAX1 is located mainly in mitochondria and acts as an anti-apoptosis factor in cells.<sup>[28]</sup> HAX1 was observed to inhibit the apoptosis of prostate cancer through the suppression of caspase-9 activation.<sup>[29]</sup> HAX1 also promoted the carcinogenesis of colorectal cancer.<sup>[30]</sup> HAX1 overexpression was observed in lung cancer.<sup>[28]</sup> ATP5MD is a component of the Fo subunit of the mitochondrial H<sup>+</sup>-ATP synthase (F-ATPase)<sup>[19]</sup> and vacuolar proton pump (V-ATPase).<sup>[31]</sup> The H<sup>+</sup>-ATP synthase and F-ATPase were identified to promote cancer progression.<sup>[32]</sup> ATP5MD is highly expressed in various cancer types.<sup>[33]</sup> Cytochrome oxidase (COX) activity was identified to be critical for the cell proliferation of lung cancer by favoring metabolic reprogramming required for cell proliferation.<sup>[34]</sup> Overexpression of COX6A1 significantly suppressed Bax- and N-(4-hydroxyphenyl) retinamide (4-HPR)-induced apoptosis in glioblastoma cell. The 4-HPR-induced mitochondrial translocation of Bax, release of mitochondrial cytochrome c, and activation of caspase-3 were attenuated under the overexpression of COX6A1 in cancer cell.<sup>[35]</sup> Current study identified HAX1, ATP5MD, and COX6A1 as hub genes and prognostic markers of KRAS MT

LUAD. Besides, those 3 genes were mitochondrial related, and our GSEA analysis also found the mitochondrial related terms were enriched in KRAS MT LUAD (Supplemental Digital Content [Table S6, <http://links.lww.com/MD/E650>]).

Our study had potential limitations. First, there was no validation performance of KRAS MT LUAD and normal tissues since the lack of KRAS MT LUAD related studies. Second, the validation of the survival hub genes in KRAS MT LUAD patients by KM PLOT was limited by the fact that there was no KRAS information in this website tool. Nevertheless, our study proved those LUAD prognostic factors might be caused by KRAS mutation.

In conclusion, our study firstly used WGCNA technique to identify KRAS MT LUAD related gene modules. The function analyses found the ribosome related terms were enriched in KRAS MT LUAD. We also identified 8 hub genes whose mRNA expression was correlated with the prognosis of KRAS MT LUAD patients. The 7 of them were validated by KM PLOT (TMSB10, RPS16, MRPL27, COX6A1, HAX1, RPL38, and ATP5MD). Further molecular biological experiments are required to explore the deep mechanism between those genes and KRAS MT LUAD.

## Author contributions

**Administrative support:** Xian Wang.

**Collection and assembly of data:** Rongkai Shi and Shuting Han.

**Conception and design:** Dongjun Dai and Xian Wang.

**Data analysis and interpretation:** Dongjun Dai and Rongkai Shi.

**Final approval of manuscript:** Dongjun Dai, Rongkai Shi,

Shuting Han, Hongchuan Jin, Xian Wang.

**Manuscript writing:** Dongjun Dai, Rongkai Shi, Shuting Han,

Hongchuan Jin, Xian Wang.

**Provision of study materials or patients:** Rongkai Shi.

## References

- [1] Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- [2] Zarogoulidis K, Zarogoulidis P, Darwiche K, et al. Treatment of non-small cell lung cancer (NSCLC). *J Thorac Dis* 2013;5(suppl):S389–96.
- [3] Meza R, Meernik C, Jeon J, et al. Lung cancer incidence trends by gender, race and histology in the United States, 1973–2010. *PLoS One* 2015;10:e0121323.
- [4] Cox AD, Der CJ. Ras history: the saga continues. *Small GTPases* 2010;1:2–7.
- [5] Adjei AA. Blocking oncogenic Ras signaling for cancer therapy. *J Natl Cancer Inst* 2001;93:1062–74.
- [6] Mao C, Qiu LX, Liao RY, et al. KRAS mutations and resistance to EGFR-TKIs treatment in patients with non-small cell lung cancer: a meta-analysis of 22 studies. *Lung Cancer* 2010;69:272–8.
- [7] Guibert N, Ilie M, Long E, et al. KRAS mutations in lung adenocarcinoma: molecular and epidemiological characteristics, methods for detection, and therapeutic strategy perspectives. *Curr Mol Med* 2015;15:418–32.
- [8] Cox AD, Fesik SW, Kimmelman AC, et al. Drugging the undruggable RAS: mission possible? *Nat Rev Drug Discov* 2014;13:828–51.
- [9] Roman M, Baraibar I, Lopez I, et al. KRAS oncogene in non-small cell lung cancer: clinical perspectives on the treatment of an old target. *Mol Cancer* 2018;17:33.
- [10] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 2008;9:559.
- [11] Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* 2013;45:1113–20.
- [12] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* 2018;68:7–30.
- [13] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16:284–7.
- [14] Majumder PK, Febbo PG, Bikoff R, et al. mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. *Nat Med* 2004;10:594–601.
- [15] Scholl C, Frohling S, Dunn IF, et al. Synthetic lethal interaction between oncogenic KRAS dependency and STK33 suppression in human cancer cells. *Cell* 2009;137:821–34.
- [16] Sun JM, Hwang DW, Ahn JS, et al. Prognostic and predictive value of KRAS mutations in advanced non-small cell lung cancer. *PLoS One* 2013;8:e64816.
- [17] Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 2005;4: Article17.
- [18] Truitt ML, Ruggero D. New frontiers in translational control of the cancer genome. *Nat Rev Cancer* 2016;16:288–304.
- [19] Meyer B, Wittig I, Trifilieff E, et al. Identification of two proteins associated with mammalian ATP synthase. *Mol Cell Proteomics* 2007;6:1690–9.
- [20] Bastide A, David A. The ribosome, (slow) beating heart of cancer (stem) cell. *Oncogenesis* 2018;7:34.
- [21] Guimaraes JC, Zavolan M. Patterns of ribosomal protein expression specify normal and malignant human cells. *Genome Biol* 2016; 17:236.
- [22] Fonseca A, Gubitoso MD, Reis MS, et al. A new approach for identification of cancer-related pathways using protein networks and genomic data. *Cancer Inform* 2015;14(suppl):139–49.
- [23] Karan D, Kelly DL, Rizzino A, et al. Expression profile of differentially-regulated genes during progression of androgen-independent growth in human prostate cancer cells. *Carcinogenesis* 2002;23:967–75.
- [24] Gentilella A, Kozma SC, Thomas G. A liaison between mTOR signaling, ribosome biogenesis and cancer. *Biochim Biophys Acta* 2015;1849: 812–20.
- [25] Chung J, Kuo CJ, Crabtree GR, et al. Rapamycin-FKBP specifically blocks growth-dependent activation of and signaling by the 70 kd S6 protein kinases. *Cell* 1992;69:1227–36.
- [26] Sribenja S, Li M, Wongkham S, et al. Advances in thymosin  $\beta$ 10 research: differential expression, molecular mechanisms, and clinical implications in cancer and other conditions. *Cancer Invest* 2009; 27:1016–22.
- [27] Li Z, Qu L, Zhong H, et al. [Mechanism of thymosin beta 10 inhibiting the apoptosis and prompting proliferation in A549 cells]. *Zhongguo Fei Ai Za Zhi* 2014;17:783–8.
- [28] Trebinska A, Rembiszewska A, Ciosek K, et al. HAX-1 overexpression, splicing and cellular localization in tumors. *BMC Cancer* 2010;10:76.
- [29] Yan J, Ma C, Cheng J, et al. HAX-1 inhibits apoptosis in prostate cancer through the suppression of caspase-9 activation. *Oncol Rep* 2015; 34:2776–81.
- [30] Li X, Jiang J, Yang R, et al. Expression of HAX-1 in colorectal cancer and its role in cancer cell growth. *Mol Med Rep* 2015;12:4071–8.
- [31] Kontro H, Hulmi JJ, Rahkila P, et al. Cellular and tissue expression of DAPIT, a phylogenetically conserved peptide. *Eur J Histochem* 2012; 56:e18.
- [32] Esparza-Molto PB, Cuezva JM. The role of mitochondrial H (+)-ATP synthase in cancer. *Front Oncol* 2018;8:53.
- [33] Kontro H, Cannino G, Rustin P, et al. DAPIT over-expression modulates glucose metabolism and cell behaviour in HEK293T cells. *PLoS One* 2015;10:e0131990.
- [34] Rak M, Benit P, Chretien D, et al. Mitochondrial cytochrome c oxidase deficiency. *Clin Sci (Lond)* 2016;130:393–407.
- [35] Eun SY, Woo IS, Jang HS, et al. Identification of cytochrome c oxidase subunit 6A1 as a suppressor of Bax-induced cell death by yeast-based functional screening. *Biochem Biophys Res Commun* 2008; 373:58–63.