Research Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2024.55.008713

Cluster Analysis of RNA m^{6}A Methylation Regulators Based on TCGA and GEO Database in Hepatocellular Carcinoma

Xiaoshi Zhang[1], Zhuo Yu[2], Xiaoyi Zhang[3] and Jianqiang Cai[1]*

[1]*National Cancer Center, National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, China*

[2]*Beijing Hospital, China*

[3]*Exchange, Development & Service Center for Science & Technology Talents, China*

***Corresponding author:** Jianqiang Cai, Department of Hepatobiliary Surgery, National Cancer Center/National Clinical Research Center for Cancer, Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College Panjiayuan, Chaoyang District, in the South, 17th, Beijing, China

ARTICLE INFO

Received: March 03, 2024

Published: March 12, 2024

Citation: Xiaoshi Zhang, Zhuo Yu, Xiaoyi Zhang and Jianqiang Cai. Cluster Analysis of RNA m^{6}A Methylation Regulators Based on TCGA and GEO Database in Hepatocellular Carcinoma. Biomed J Sci & Tech Res 55(3)-2024. BJSTR. MS.ID.008713.

ABSTRACT

As potential targets and candidate molecules for the treatment of cancer, m^{6}A RNA and its regulator genes have important application prospects in individualized therapy of Hepatocellular Carcinoma(HCC). Through TCGA and GEO data screening, we demonstrated the expression of different m^{6}A regulators within HCC related to genetic variation, biological pathways and patients' prognosis. We use machine learning methods to cluster m^{6}A regulators and differential expressed genes, which may provide a theoretical basis for the establishment of a risk assessment model.

Keywords: Hepatocellular Carcinoma; m^{6}A RNA; Database; Prognosis

Abbreviations: HCC: Hepatocellular Carcinoma; ICGC: International Cancer Genome Consortium; GEO: Gene Expression Omnibus; GO: Gene Ontology; MF: Molecular Function; BP: Biological Pathway; CC: Cellular Component; DEGs: Differential Expression Genes; SNV: Single Nucleotide Variation; KEGG: Kyoto Encyclopedia of Genes and Genomes; LASSO: Least Absolute Shrinkage and Selection Operator; OS: Overall Survival; GSVA: Gene Set Variation Analysis

Introduction

In recent years, as one of the most common and fatal malignant tumors in the world, liver cancer has become the focus of medical and public attention. According to the statistics of the World Health Organization, the incidence of liver cancer ranks fifth and the mortality ranks third, causing more than 800,000 deaths each year, and the number is still increasing [1-4]. As the main pathological type, more than 90% of liver cancer is hepatocellular carcinoma. Unfortunately, because the onset of HCC is hidden and there are no obvious symptoms in its early stage, most patients are found in the middle and late stage, which causes poor therapeutic effect [5]. Although great progress has been made in targeting, immunity, intervention, radiotherapy and other fields in recent years, the prognosis of patients with HCC is still not ideal enough [6]. Therefore, the study of HCC is still very urgent and of great significance. m6A RNA refers to the methylation of the sixth nitrogen atom of adenine in RNA. Since it was first discovered in the 1970s, m6A methylation has gradually become a research hotspot in life science [7]. M^{6}A RNA can regulate many biological processes such as RNA degradation, transcription, splicing and translation, thus affecting gene expression and protein synthesis, and plays an important regulatory role in the occurrence and development of HCC. In HCC, the increase of m^{6}A RNA expression is mainly regulated by three key proteins: «writer», «eraser» and «reader». More and more studies have shown that the abnormal expression and regulatory mechanism of m^{6}A RNA are closely related to the occurrence, progression and prognosis of HCC [8-10].

As potential targets and candidate molecules for the treatment of liver cancer, m^{6}A RNA and its related genes have important application prospects in individualized therapy and immunotherapy of HCC. The purpose of this study is to use bioinformatics technology to analyze the gene chips of patients with HCC in public database, comprehensively and systematically sort out and analyze the mutation, expression and survival differences of m^{6}A RNA and its related genes in HCC. Through the study of the expression level, function, regulatory mechanism, related proteins and signal pathways of m^{6}A RNA, we can better understand the key role of m^{6}A RNA in the occurrence, progression and prognosis of HCC. Through a comprehensive and in-depth exploration of the study of m^{6}A RNA in HCC, we are expected to provide new targets and methods for early diagnosis, treatment and prognosis evaluation, provide more effective and individual treatment for patients, and contribute to the development and progress of liver cancer research.

Materials and Methods

Data Selection and Processing

The RNA-Seq gene expression profiles of patients with HCC were downloaded from The Cancer Genome Atlas(TCGA) portal (https://cancergenome.nih.gov/) and International Cancer Genome Consortium (ICGC) portal (https://dcc.icgc.org/). We downloaded gene chip GSE10143 from NCBI Gene Expression Omnibus(GEO) database (http://www.ncbi.nlm.nih.gov/geo/). All the samples used by GSE10143 were resected from patients with HCC, including 80 tumor tissue samples and 82 normal liver tissue samples. The platform is GPL5474, Human 6k Transcriptionally Informative Gene Panel for DASL.

Differential Expression Gene Analysis

Use the "Limma" package of R software (version 4.1.2) to analyze the differential expression of mRNA in the dataset. The data is selected as adjusted «P < 0.05 and |logFC|> 2» (logFC is defined as threshold mRNA differential expression screening) as Differential Expression Genes(DEGs). The heat map was drawn by R software package "pheat" thermal map.

Functional Enrichment Analysis

In order to further screen and confirm the potential function of DEGs, the functional enrichment analysis of the data was carried out by means of Gene Ontology (GO). GO is a widely used tool to annotate genes with potential functions through Molecular Function (MF), Biological Pathway (BP), and Cellular Component (CC). Kyoto gene and genome database (Kyoto Encyclopedia of Genes and Genomes, KEGG) enrichment analysis is a method used to annotate gene function and related signal pathways. In this study, the Cluster function enrichment and KEGG pathway of DEGs were analyzed by using the "Cluster bubble Profiler" package in R language software, and the bubble diagram was drawn by R software package "ggord".

Establishment of m^{6}A Subgroups

All genes were used to conduct a univariate Cox survival regression. P<0.1 was used to create a Least Absolute Shrinkage and Selection Operator (LASSO) regression model. R package "glmnet", "forest" and "survival" were used during this process. Unsupervised clustering was performed to identify subgroups based on the m^{6}A regulators. Package "Consensus Cluster Plus" in R software were used.

Statistical Analysis

Categorical data were analyzed by chi-square test or Fisher exact test. Normal data were analyzed using the t-test, and non-normal data were analyzed using the Wilcox on rank sum test. The Kaplan-Meier method was used for survival analysis. The COX model was used for univariate and multifactorial analyses. Differences were considered statistically significant at $P < 0.05$. All results were double counted three times.

Results

Variation of m^{6}A Regulators in HCC

By downloading data from TCGA and GEO databases, we collected 595 liver cancer patients, including 480 HCC samples and 115 normal samples. Copy number variations including both gain and loss of copies were found in 23 m^{6}A regulated genes (Figure 1A). The position of the CNV change of the 23 m^{6}A regulated genes on the chromosome is shown in the (Figure 1B). Considered as DEGs by adjusted P<0.05, differences in expression were found in 21 of all 23 m^{6}A regulated genes (Figure 1C). The gene mutation data are also downloaded from the TCGA database, among which 364 samples have complete mutation data. Single Nucleotide Variation (SNV) were found in 31 samples including synonymous mutation, missense mutation, shift mutation and nonsense mutation , accounting for 8.52% of the total (Figure 1D).

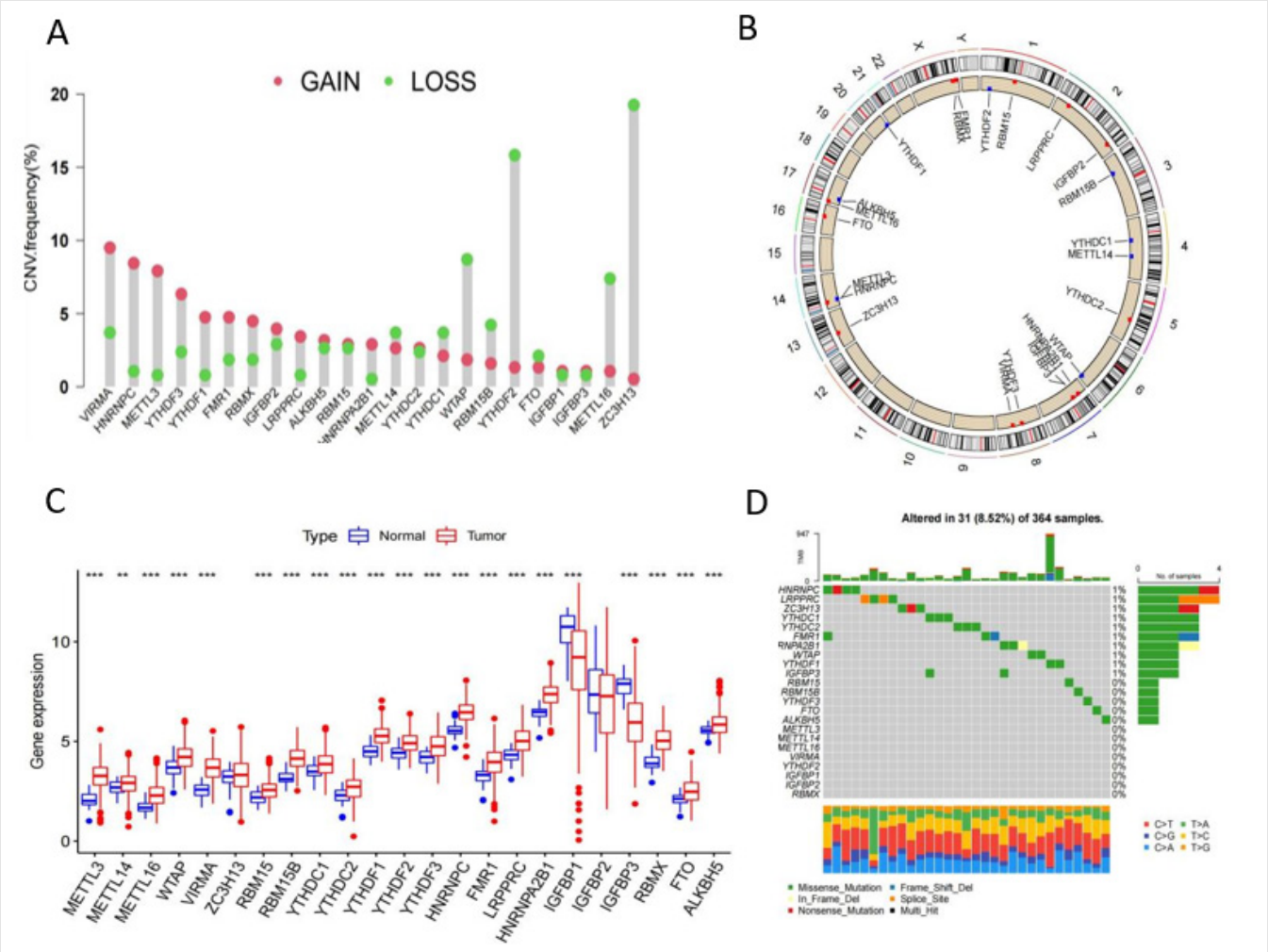

Figure 1: Landscape of genetic and expression variation of m^{6}A regulators in HCC.

A. Landscape of CNV in 23 m^{6}A regulated genes in TCGA-LIHC.

B. The location of CNV alteration of m^{6}A regulators on 23 chromosomes using TCGA-LIHC.

C. Boxplot to intuitively illustrate the differences between tumor and normal tissues in TCGA-LIHC.

D. Landscape of single nucleotide variation in TCGA-LIHC.

Survival Analysis of m^{6}A Regulators in HCC

The survival data of HCC patients were downloaded from TCGA database and analyzed by Kaplan-Meier survival analysis. The statistical difference of Overall Survival (OS) was distinguished by $P < 0.05$. A total of 6 of the 23 m^{6}A regulated genes can affect the prognosis of patients with HCC (Figures 2A-2F). High expression level all of these 6 genes promote the development of HCC and lead to a worse prognosis.

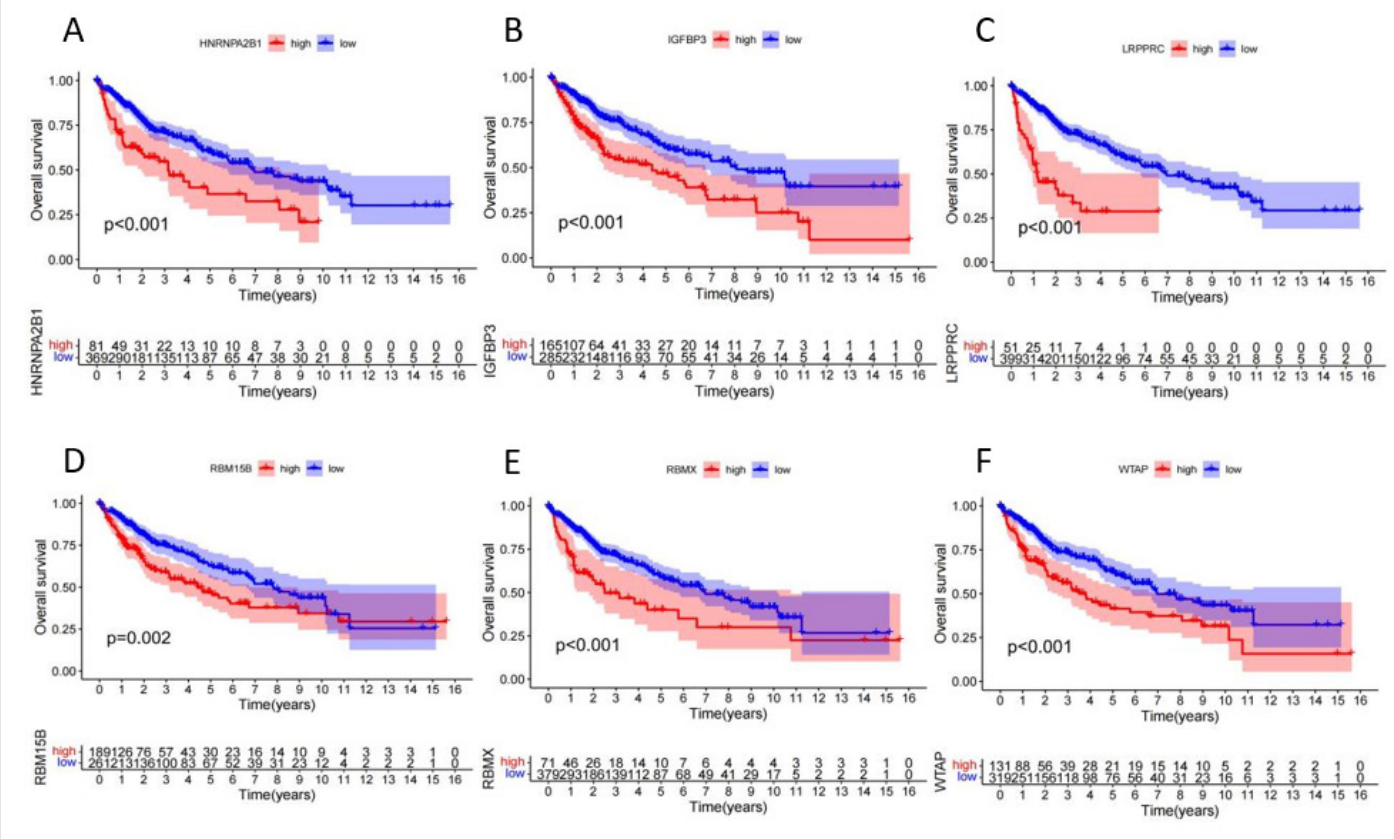

Figure 2: Kaplan-Meier curve of m^{6}A regulated genes in TCGA-LIHC.

A. High expression of HNRNPA2B1 is associated with poor prognosis in patients with HCC.
B. High expression of IGFBP3 is associated with poor prognosis in patients with HCC.
C. High expression of LRPPRC is associated with poor prognosis in patients with HCC.
D. High expression of RBM15B is associated with poor prognosis in patients with HCC.
E. High expression of RBMX is associated with poor prognosis in patients with HCC.
F. High expression of WTAP is associated with poor prognosis in patients with HCC.

Cluster Analysis of m^{6}A RNA in HCC

According to the difference of m^{6}A regulated gene expression, we clustered all the HCC samples obtained from TCGA database. This process is accomplished through machine learning. We tried to divide all the samples into 2-9 categories, and there was the most obvious difference when they were divided into 3 categories (Figures 3A-3C). We named these three clusters as m^{6}A cluster A, B and C. Gene Set Variation Analysis (GSVA) was used to find out pathways that genes in different m^{6}A clusters enriched in. The results show that genes are enriched in cell homeostasis, cellular micro-environment, metabolism and a variety of tumor-related signaling pathways (Figures 4A-4C).

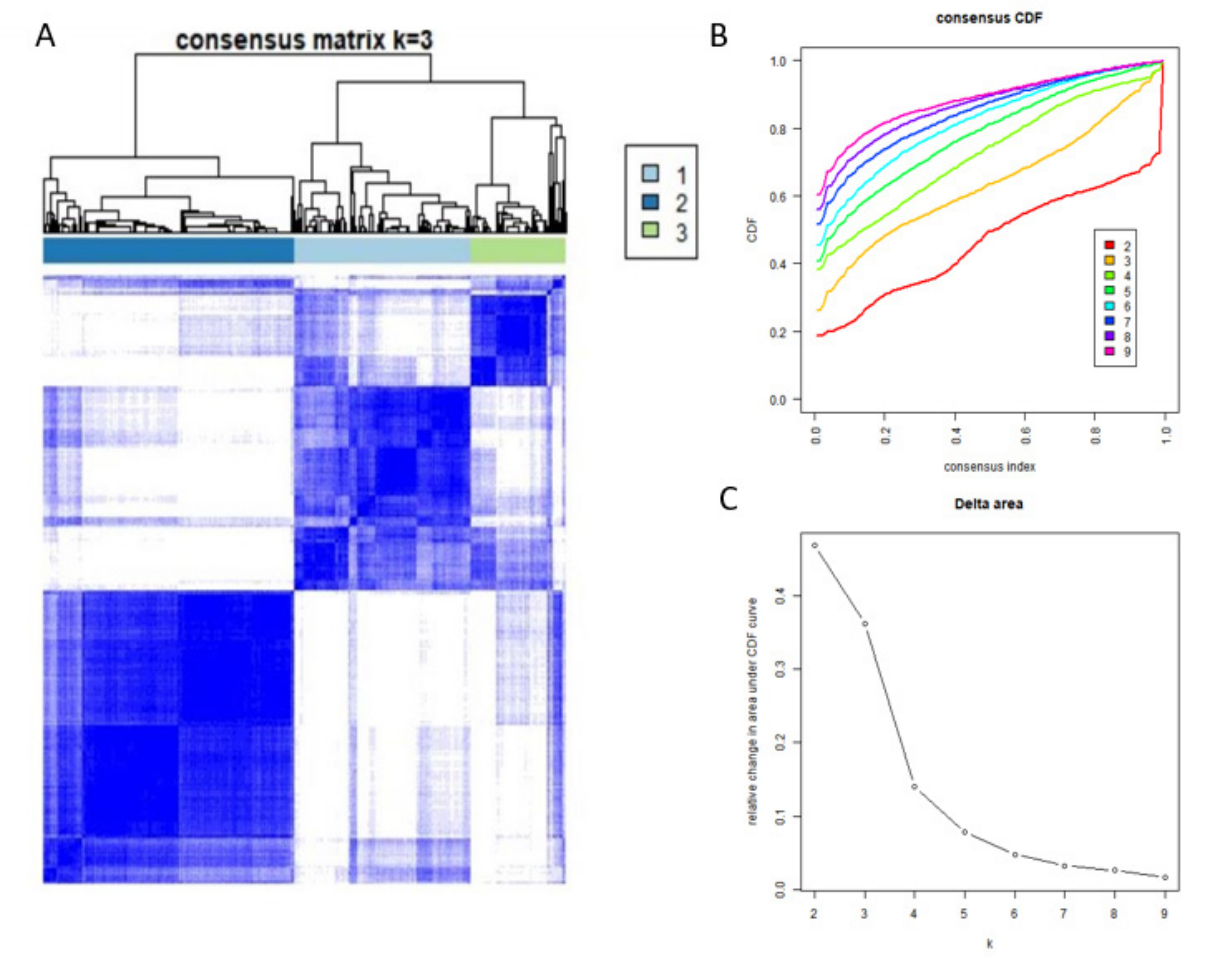

Figure 3: Unsupervised clustering of 23 m^{6}A regulators in the TCGA-LIHC cohort.

A. Colour-coded heatmap corresponding to the consensus matrix for k=3 obtained by applying consensus clustering. Colour gradients represent consensus values from 0 to 1; white corresponds to 0 and dark blue to 1.

B. Criteria for selecting number of categories. CDF line shows remarkable difference between different clusters.

C. Criteria for selecting number of categories. Delta area line shows remarkable difference between different clusters.

DOI: 10.26717/BJSTR.2024.55.008713

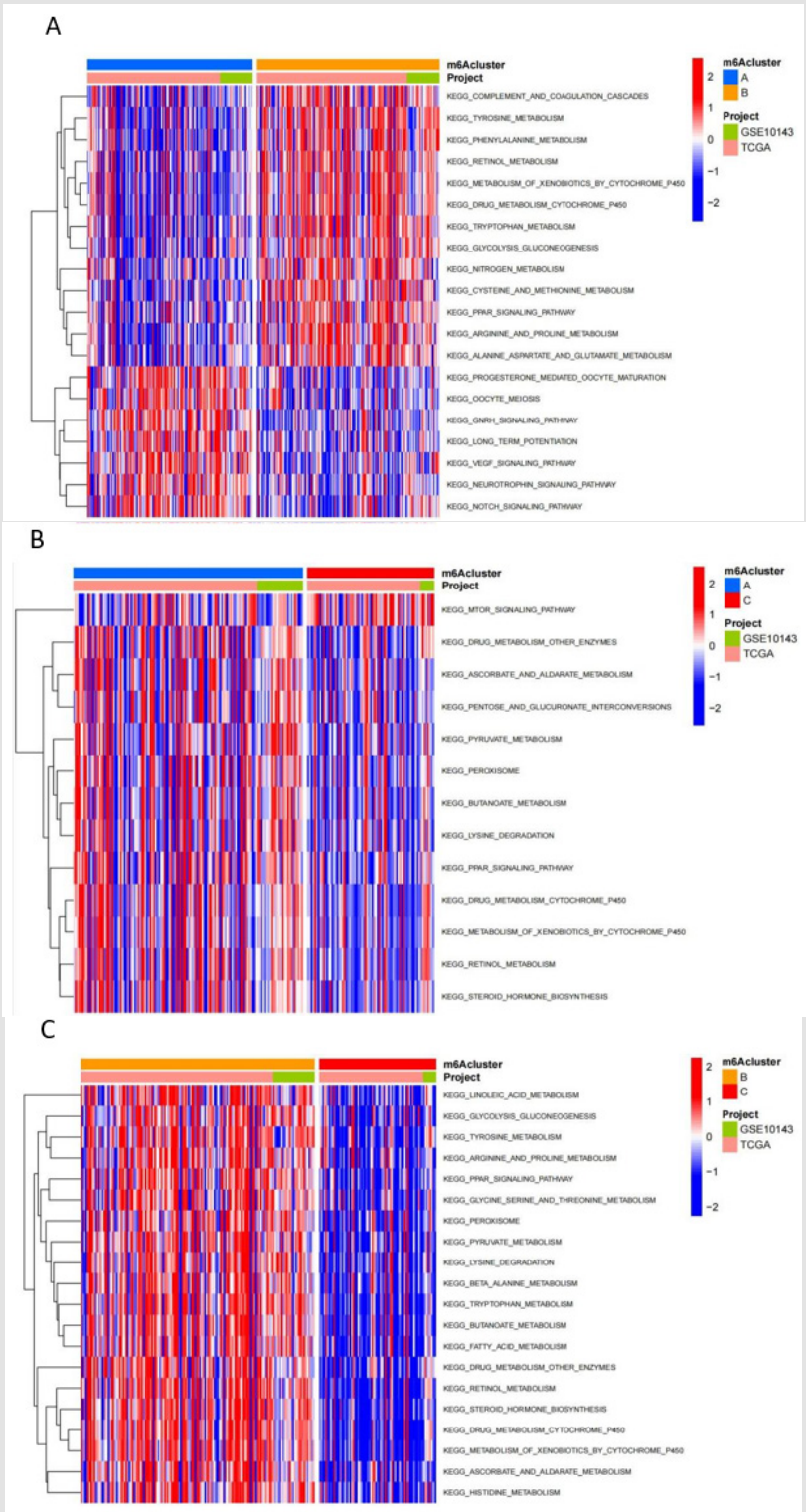

Figure 4: M^{6}A gene regulation in pathways landscape of HCC in TCGA and GEO database.

A. Pathways of m6A gene regulation enrichment between m^{6}A cluster A and B.

B. Pathways of m6A gene regulation enrichment between m^{6}A cluster A and C.

C. Pathways of m6A gene regulation enrichment between m^{6}A cluster B and C.

Cluster Analysis of Differential Genes by m^{6}A Clusters in HCC

There are also differentially expressed genes in different m^{6}A clusters. In order to find out these genes, we found the gene expression data of individuals in different m^{6}A clusters from the TCGA database. Then, we clustered these genes using the method mentioned above. After we divided all the samples into 2-9 categories, there was the most obvious difference when they were divided into 3 categories (Figure 5A). We named these three clusters as gene cluster A, B and C. A total of 1046 genes show difference between gene cluster A and B. The number of genes between cluster A and C is 94 while the number between cluster B and C is 2109 (Figure 5B). To explore the biological functions of these genes, they were categorized into BP, CC and MF. Under stringent threshold conditions by adjusted P value less than 0.05, top 10 terms of each part are in (Figure 5C). Through Kaplan-Meier survival analysis, we found that there were differences in individual prognosis among different gene clusters. The overall survival time of patients in gene cluster A was the longest while in cluster C was the shortest (Figure 5D). The basic information, general situation, tumor stage and other data of these individuals are shown in the (Figure 5E). It is worth mentioning that we analyzed the differentially expressed genes that affect the prognosis of HCC patients in different gene clusters and found that m^{6}A regulated genes were all differentially expressed in different gene clusters (Figure 6). However, it needs further research.

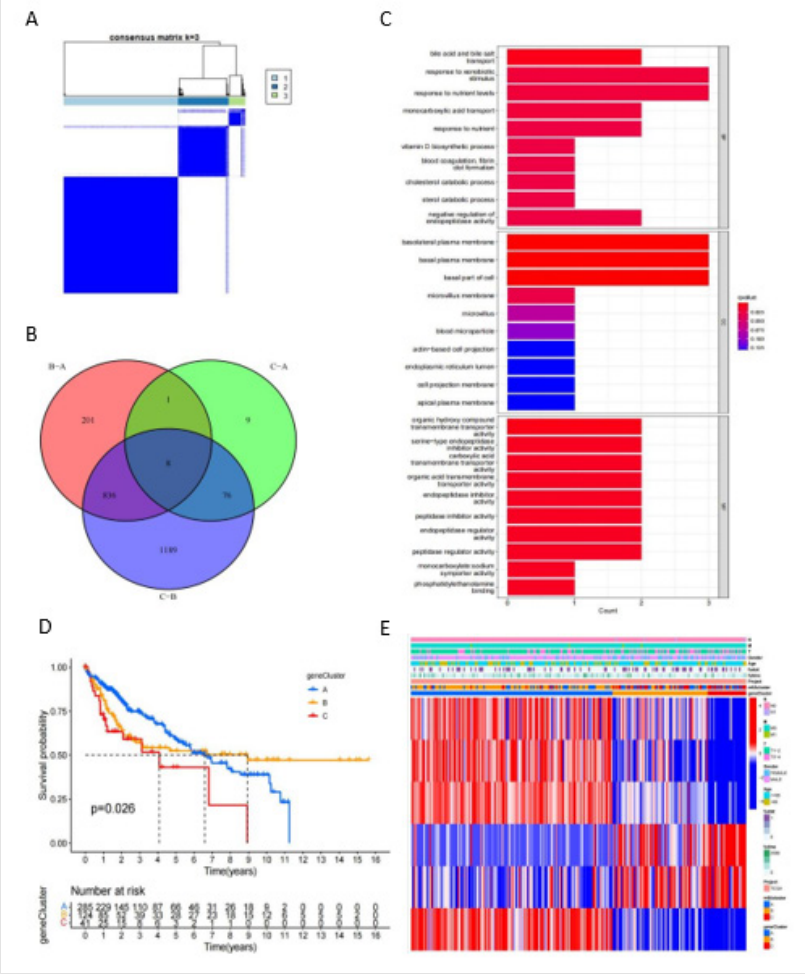

Figure 5: Landscape of differential genes by m^{6}A clusters.

A. Unsupervised clustering of differential genes by m^{6}A clusters.
B. Venn diagram of genes between different m^{6}A clusters.
C. Bar-plot of GO enrichment in cellular component terms, biological process terms and molecular function terms.
D. Kaplan-Meier curve of different gene clusters.
E. Landscape of basic information, general situation, tumor stage of individuals in different m6A clusters and gene clusters.

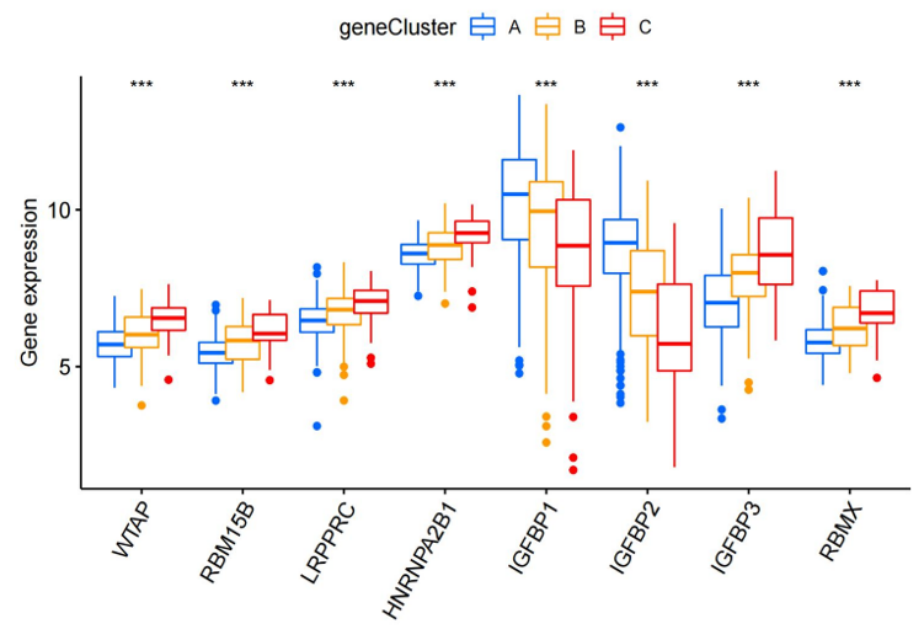

Figure 6: Expression of m^{6}A regulated genes in different gene clusters.

Discussion

In recent years, with the deepening of research, the important role of m^{6}A modification in HCC has been widely recognized. More and more studies have gradually revealed the process of m^{6}A modification and the effect of m^{6}A regulators on the biological behavior of HCC [11,12]. For example, Chen, et al. [13]) found that METTL3 in human HCC was significantly u up regulated compared with non-tumor liver control, and the overall m6A modification level in human HCC was also higher than that in normal individuals. At the same time, the team also proved that METTL3-mediated m^{6}A hypermethylation is a new mechanism of epigenetic silencing of tumor suppressor gene expression in human cancer. Another study also proved that m^{6}A modification is very important for the regulation of EMT process of HCC. The level of m^{6}A modification increases significantly when EMT occurs, which can affect the invasion, metastasis and EMT process of HCC both in vivo and in vitro [14]. (Lin, et al. [15]) further identified Snail as an important transcription factor involved in EMT, has become the target of METTL3-mediated m^{6}A modification. METTL3 works with YTHDF1 to promote protein translation of Snail, explaining how overexpression of METTL3 promotes HCC transfer. More and more m^{6}A regulated genes have been found. For example, Chen, et al. [16,17]) found that the other two components of m^{6}A complex, WTAP and KIAA1429, were significantly up regulated in HCC, which was related to the low overall survival rate of patients (Ma, et al. [18]).

Proved that METTL14 interacts with microprocessor protein DGC8 to promote the maturation of microRNA-126, and the downregulation of METT14 weakens the expression of microRNA-126, thus promoting HCC transfer (Li, et al.[19]). Observed the overexpression of FTO in HCC tissues. At the same time, with the increase of m^{6}A level, the knockdown of FTO induced cell cycle arrest and inhibited the colony formation ability of HCC cells. In addition, it has been reported that the expression of YTHDF2 is down-regulated in human HCC, and the loss of YTHDF2 destroys the m^{6}A-dependent mRNA decay of IL11 and SERPINE2, resulting in increased invasiveness of HCC [20]. It can be seen that m^{6}A RNA is closely related to HCC. It has been confirmed that different gene mutations may be the key factors leading to the change of m^{6}A RNA expression level. In addition, some m^{6}A regulator genes have also been confirmed to play a cancer-promoting role in a variety of malignant tumors [21-23], such as non-small cell lung cancer [24], gastric cancer [25], bladder cancer [26], colon cancer [27] and so on. More significantly, it has also been confirmed to be associated with drug resistance of cancer, which indirectly leads to a poor prognosis of tumor patients [28]. For this reason, we analyzed the large sample size data in TCGA and GEO database to find the differentially expressed m^{6}A regulator gene which may affect the prognosis of HCC patients. Through LASSO regression and other machine learning algorithms, we carried out cluster analysis of these differentially expressed genes, and enriched and analyzed the biological pathways that each type of genes may affect.

Then, we analyze the differences of gene expression in different m^{6}A clusters. We clustered the different genes by machine learning method and analyzed the clinical manifestations of individuals in different clusters. By detecting the expression of these genes, we can establish a prognostic risk model for patients with HCC, which may

be helpful for the diagnosis and treatment of HCC patients. However, it needs to be verified by further experiments and clinical cohort studies. Our study had some limitations. First, although we found that m^{6}A regulators play different roles in the alternative pathways, the potential molecular mechanisms were not evaluated, it warrants further investigation. Second, m^{6}A clusters and their risk gene clusters might improve prognosis of HCC patients. Further research is necessary to explore whether these genes could be used as diagnostic markers or therapeutic targets in HCC and guiding more effective treating strategies.

Conclusion

In summary, m^{6}A regulators have CNV and SNV mutations in HCC patients, which may lead to poor prognosis of HCC. Different clusters of m^{6}A regulators play different roles in multiple biological pathways. Different gene clusters can affect the prognosis of patients with HCC. There is a differential expression of m^{6}A regulators in different gene clusters. The comprehensive evaluation of m^{6}A modification pattern in HCC will contribute to enhancing our understanding of tumor characterization and may guide more effective therapy of HCC patients.

Author Contributions

Xiaoshi Zhang designed this study, Zhuo Yu collected the data, Xiaoyi Zhang and Zhuo Yu analysised the data, Xiaoshi Zhang wrote this manuscript. Jianqiang Cai revised the manuscript. All authors read and approved the final manuscript.

Conflict of Interest Statement

There was no potential conflicts of interest to declare among the authors.

Funding Sources

This work was supported by the CAMS Innovation Fund for Medical Sciences (CIFMS) (grant number 2021-I2M-1-066) and the Sanming Project of Medicine in Shenzhen (grant number SZSM202011010).

Data Availability Statement

The datasets generated from this research can be disclosed only in specific circumstances. Further inquiries can be directed to the corresponding author.

References

1. (2018) European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol 69: 182-236.
2. (2015) GBD 2013 Mortality, Causes of Death Collaborators. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the Global Burden of Disease Study 2013. Lancet 385: 117-171
3. Han B, Zheng R, Zeng H, Shaoming Wang, Kexin Sun, et al. (2024) Cancer incidence and mortality in China, 2022. J Natl Cancer Cent.
4. Wu C, Li M, Meng H, Yukun Liu, Weihong Niu, et al. (2019) Analysis of status and countermeasures of cancer incidence and mortality in China. Sci China Life Sci 62(5): 640-947.
5. Vogel A, Cervantes A, Chau I, B Daniele, JM Llovet, et al. (2018) Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 29(Suppl 4): iv238-iv255.
6. European Association for The Study of The Liver, European Organization for Research and Treatment of Cancer. (2012) EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma. J Hepatol 56(4): 908-943.
7. Perry J, Kelley D (1994) Existence of methylated messenger RNA inmouse L cells. Cell 1(1): 37-42.
8. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Mali Salmon-Divon, Lior Ungar, et al. (2012) Topology of the human and mouse m^{6}A RNA methylomes revealed by m^{6}A-seq. Nature 485(7397): 201-206.
9. Tong J, Flavell RA, Li HB (2018) RNA m^{6}A modification and its function in diseases. Front Med 12(4): 481-489.
10. Han D, Liu J, Chen C, Lihui Dong, Yi Liu, et al. (2019) Anti-tumour immunity controlled through mRNA m^{6}A methylation and YTHDF1 in dendritic cells. Nature 566(7743): 270-274.
11. He L, Li H, Wu A, Yulong Peng, Guang Shu, et al. (2019) Functions of N6-methyladenosine and its role in cancer. Mol Cancer 18(1): 176.
12. Dai D, Wang H, Zhu L, Hongchuan Jin Xian Wang (2018) N6-methyladenosine links RNA metabolism to cancer progression. Cell Death Dis 9(2): 124.
13. Chen M, Wei L, Law CT, Tsang FH, Shen J, et al. (2018) RNA N6-methyladenosine methyltransferase METTL3 promotes liver cancer progression through YTHDF2 dependent posttranscriptional silencing of SOCS2. Hepatology 67(6): 2254-2270.
14. Wu R, Liu Y, Zhao Y, Bi Z, Yao Y, et al. (2019) m^{6}a methylation controls pluripotency of porcine induced pluripotent stem cells by targeting SOCS3/JAK2/STAT3 pathway in a YTHDF1/YTHDF2-orchestrated manner. Cell Death Dis 10: 171.
15. Lin X, Chai G, Wu Y, Li J, Chen F, et al. (2019) RNA m6A methylation regulates the epithelial mesenchymal transition of cancer cells and translation of snail. Nat Commun 10: 2065.
16. Chen Y, Peng C, Chen J, Chen D, Yang B, et al. (2019) WTAP facilitates progression of hepatocellular carcinoma via m^{6}AHuR-dependent epigenetic silencing of ETS_1. Mol Cancer 18: 127.
17. Cheng X, Li M, Rao X, Zhang W, Li X, et al. (2019) KIAA1429 regulates the migration and invasion of hepatocellular carcinoma by altering m6A modification of ID_2 mRNA. Onco Targets Ther 12: 3421-3428.
18. Ma JZ, Yang F, Zhou CC, Liu F, Yuan JH, et al. (2017) METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N6 -methyladenosine-dependent primary MicroRNA processing. Hepatology 65(2): 529-543.
19. Li J, Zhu L, Shi Y, Liu J, Lin L, et al. (2019) m6A demethylase FTO promotes hepatocellular carcinoma tumorigenesis via mediating PKM2 demethylation. Am J Transl Res 11: 6084-6092.
20. Hou J, Zhang H, Liu J, Zhao Z, Wang J, et al. (2019) YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. Mol Cancer 18: 163.
21. Li H, Han X, Song L, Liwei Zhang, Zhibo Jin, et al. (2024) LINC00645 inhib-

its renal cell carcinoma progression by interacting with HNRNPA2B1 to regulate the ROCK1 mRNA stability. Gene 5(905): 148232.

22. Wang X, Zhi M, Zhao W, Jiayin Deng (2024) HNRNPA2B1 promotes oral squamous cell carcinogenesis via m^6 A-dependent stabilization of FOXQ1 mRNA stability. IUBMB Life.
23. Zhang J, Liu B, Xu C, Chenchen Ji, Anan Yin, et al. (2023) Cholesterol homeostasis confers glioma malignancy triggered by hnRNPA2B1-dependent regulation of SREBP2 and LDLR. Neuro Oncol 9: 233.
24. Yao Y, Chen C, Wang J, Libao Gong, Wen Gao, et al. (2023) Circular RNA circATP9A promotes non-small cell lung cancer progression by interacting with HuR and by promoting extracellular vesicles-mediated macrophage M_2 polarization. J Exp Clin Cancer Res 42(1): 330.
25. Chen H, Xiao L, Xie G, Peng Zhang, Ping Dong, et al. (2024) LINC00355 promotes gastric carcinogenesis by scaffolding p300 to activate CDC42 transcription and enhancing HNRNPA2B1 to stabilize CDC42 mRNA dependent on m6A. Mol Carcinog 63(3): 430-447.
26. Zhao Y, Chen J, Zheng H, Tianxin Lin, Changhao Chen, et al. (2024) SUMOylation-Driven mRNA Circularization Enhances Translation and Promotes Lymphatic Metastasis of Bladder Cancer. Cancer Res 84(3): 434-448.
27. Lu Y, Zou R, Gu Q, Xinyue Wang, Junying Zhang, et al. (2023) CRNDE mediated hnRNPA2B1 stability facilitates nuclear export and translation of KRAS in colorectal cancer. Cell Death Dis 14(9): 611.
28. Jiang J, Zhu J, Qiu P, Jie Ni, Wei Zhu, et al. (2023) HNRNPA2B1-mediated m6A modification of FOXM1 promotes drug resistance and inhibits ferroptosis in endometrial cancer via regulation of LCN2. Funct Integr Genomics 24(1): 3.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2024.55.008713

Jianqiang Cai. Biomed J Sci & Tech Res

Submission Link: https://biomedres.us/submit-manuscript.php