Original Research

Gene signature and prognostic merit of M6a regulators in colorectal cancer

Jinfeng Zhang[1,*], Xuedi Cheng[2,*], Junzheng Wang[3], Yongjie Huang[4], Junhui Yuan[5] and Dawen Guo[1]

[1]Department of Clinical Laboratory, First Affiliated Hospital of Harbin Medical University, Harbin City, Heilongjiang Province 150001, China; [2]Department of Clinical Laboratory, Affiliated Hospital of Qingdao University, Qingdao City, Shandong Province 266000, China; [3]Department of Stomatology, Qingdao Haici Medical Group, Qingdao City, Shandong Province 266034, China; [4]Department of General Surgery, Hebei Provincial Hospital of Traditional Chinese Medicine, Shijiazhuang City, Hebei Province 050011, China; [5]Department of Breast &Thyroid Surgery, Qingdao Women and Children's Hospital, Qingdao City, Shandong Province 266034, China
Corresponding authors: Junhui Yuan. Email: junhuiyuan111@163.com; Dawen Guo. Email: professorguo618@163.com
*These authors contributed equally to this paper.

Impact statement

Although new diagnostic techniques and treatments are increasingly updated for CRC, the clinical outcomes of CRC patients are still not encouraging with a low survival rate. N6-methyladenosine (m6A) as a popular modification on mRNA is associated with multiple types of cancers. Our purpose is to identify gene signature and prognostic ability of m6A modulators in CRC. For the first time, we identified genetic changes of m6A modulators and built prognostic gene signature in CRC, which may provide effective targets for the diagnosis and management of CRC.

Abstract

Although new diagnostic techniques and treatments are increasingly updated, the clinical outcomes of CRC patients are still not encouraging with a low survival rate. N6-methyladenosine as a popular modification on mRNA is associated with multiple types of cancers. Our purpose is to evaluate gene signature and prognostic ability of N6-methyladenosine in CRC. We used the gene expression, copy number variation, simple nucleotide variation and clinical messages from The Cancer Genome Atlas database. We first identified mutation and copy number variations of N6-methyladenosine regulatory genes in both colon adenocarcinoma and rectum adenocarcinoma. Fourteen of all 17 N6-methyladenosine regulatory genes were related with higher mRNA expression, whereas deletion leads to reduced expression. Using univariate Cox regression analysis, RBM15, YTHDC2, and METTL14 genes in the rectum adenocarcinoma samples were conspicuously associated with the prognosis of patients. Based on the least absolute shrinkage and selection operator regression models, we built a 2-gene (YTHDC2 and IGF2BP3) signature of N6-methyladenosine regulators with prognostic ability. The 1-, 3-, and 5-year AUCs of this signature were all greater than 0.6, and the *P*-value for risk prediction for patients was also less than 0.0001. Moreover, high IGF2BP3 gene expression was significantly associated with IFN-γ in colon adenocarcinoma , and related to the azurophil granule membrane pathway in rectum adenocarcinoma. High YTHDC2 expression in colon adenocarcinoma is closely related to cell energy metabolism. In the rectum adenocarcinoma, high YTHDC2 gene expression is related to the cell centrosome pathway. In conclusion, for the first time, we identified genetic changes of N6-methyladenosine modulators and built a prognostic gene signature in CRC.

Keywords: Colorectal cancer, N6-methyladenosine RNA methylation, prognostic signature, TCGA, bioinformatics

***Experimental Biology and Medicine* 2020; 245: 1344–1354. DOI: 10.1177/1535370220936145**

Introduction

Colorectal cancer (CRC) is diagnosed as the third popular cancer in men and the second in woman worldwide.[1] CRC can be divided into two major types including COAD and READ. It is reported as the third highest prevalent cancer, accounting for roughly 10% of new cancer incidence, and the second ranked most popular reason of cancer-related mortality.[2] Although new diagnostic techniques as well as treatment methods are increasingly updated, the clinical

outcomes of the CRC patients are still not encouraging with a low rate of five years' survival because they are usually diagnosed at an advanced stage.[3] Therefore, a deep understanding of molecular biology of CRC can lead to the identification of diagnostic, prognostic, and therapeutic biomarkers.

RNA methylation modification comprised over 60% of all RNA modifications, and m6A is the most popular modification on mRNA in the majority of eukaryotes. The m6A modification happens mainly in adenine in the RRACH order, and its function is confirmed by "Writer", "Eraser," and "Reader".[4] This kind of modification has been shown to play essential functions in mRNA metabolism and diverse biological pathways,[5] especially in cancers.[6] In addition, there is increasing evidence that abnormal regulation of m6A is connected to a variety of cancers.[7] For instance, Hou *et al.*[8] characterized the m6A-mRNA landscape in human hepatocellular carcinoma. They found that YTHDF2 transcription ends at hypoxia-inducible factor 2α with important roles in transcriptome and cancer progression. In CRC study, upregulated METTL3 can promote metastasis by miR-1246/SPRED2/MAPK signaling pathway.[9] Besides, Deng *et al.*[10] found that METTL3 restrained CRC development through p38/ERK pathways.

Recently, there are increasing research reports on gene prognosis assessment signatures with the help of microarray and RNA-sequencing data. Based on the gene expression profiles of cancers, people identified various prognostic signatures in different cancers. For examples, people established a newly developed four gene-signature with predictive utility in CRC.[11] In another study, Wang *et al.*[12] used TCGA database to find a 12-gene expression signature to measure prognosis for CRC. Besides, Zuo *et al.*[13] also employed the gene expression profiles of TCGA to determine a 6-gene signature in CRC.

By literature searching, we found some papers about signature of m6A regulators in human cancers. In the head and neck squamous cell carcinoma, a two-gene marker including YTHDC2 and HNRNPC was build and may suggest patients' survival.[14] Besides, Zhou *et al.*[15] determined genetic alterations of m6A modulators in clear cell renal cell carcinoma and reveal a meaningful correlation among the alterations and poor clinical characteristics. However, there were no studies about prognostic signature based on m6A regulator genes in CRC.

For the first time, we defined genetic changes of m6A modulators in CRC. We used the gene expression, CNV, SNV, and clinical data from TCGA database. Then, 14 of all 17 m6A regulatory genes was correlated with elevated mRNA content, whereas deletion resulted in decreased expression. Using univariate Cox regression analysis, RBM15, YTHDC2, and METTL14 genes in the READ samples were notably associated with the prognosis of patients. Based on the LASSO regression models, we built a 2-gene (YTHDC2 and IGF2BP3) signature of m6A regulators with prognostic value in CRC. Moreover, high IGF2BP3 gene expression was significantly associated with IFN-γ in COAD, and related to the azurophil granule membrane pathway in READ.

Materials and methods

Datasets acquisition

The COAD and READ datasets including CNV, SNV, and mRNA expression data as well as all corresponding clinical messages used in the work were exported from the TCGA dataset (http://gdc-portal.nci.nih.gov/)[16] and downloaded in July 2019. As for CNV data download, we used the RTCGA R package (https://rtcga.github.io/RTCGA/index.html) as the level 3 files. The SNV data and the CNV download method are described above, the files were called by the mutect2 software. [17]

Data processing

For transcriptome data, we obtained a total of 456 COAD and 166 READ samples, and the download data were FPKM (Fragments Per Kilobase of exon model per Million mapped fragments) files, and were converted into TPM (Transcripts Per Kilobase of exonmodel per Million mapped reads) files. According to SNV messages, we obtained a total of 399 COAD and 137 READ samples, and the download data were level 3 after mutect processing. The R package maftools were used to read the mutect results, intron interval and the mutation annotated as silent was removed, and further the mutation characteristics of m6A gene were extracted. For CNV data, 452 COAD and 165 READ samples were as level 3 files. The "Segment_Mean" value was already included in the data list, and this value can be used as a basis for judging whether CNV has occurred. Segment_mean values smaller than -0.3 are categorized as a "loss", and Segment_mean values larger than 0.3 are categorized as a "gain". We calculate the frequency of loss and gain of each m6A gene in each sample as frequency CNV. When the frequency CNV is more than 60%, it is considered as high frequency CNV; when the frequency CNV is less than 40%, it is considered as low frequency CNV. Besides, there were 459 COAD and 170 READ samples with clinical information. After integration of all data, samples with incomplete clinical information that had a survival time of less than 30 days were excluded, and all samples had complete CNV, SNV, and mRNA data for the m6A regulators. The final samples available for survival analysis were 402 COAD and 145 READ samples.

LASSO regression model

The LASSO model is a compressed estimate. It obtains a more detailed model by building a penalty function. Thus, the superiority of subset shrinkage is retained, and it is a method for processing biased estimates with complex collinearity data. This model was generated by glmnet R package.[18]

Gene set enrichment analysis

GSEA is a calculation approach for determining whether a set of genes defined a priori show a statistically important and consistent variations among biological conditions. It was performed with available software and downloaded from the website (http://software.broadinstitute.org).[19]

All COAD and READ samples were stratified into two genotypes based on the median expression profiling level. P-value < 0.05 and a false discovery rate (FDR) < 0.25 were regarded as notable.

Statistical analysis

All statistical analyses were done in R language (version 3.6.2). The univariate Cox regression analysis was conducted to explore correlation between clinical features and CNV and/or SNV. We analyzed the patients' survival through Kaplan–Meier and log-rank test. The Jordan index was employed as the cutoff to categorize the high-and low-risk prognostic samples. Statistical experiments were two-sided in which P-value < 0.05 was thought meaningful.

Results

Mutation and CNVs of m6A regulatory genes

In the SNV mutation data of 399 COAD patients, the mutations of m6A regulatory gene appeared in 113 independent samples (Supplementary Table 1). Among them, the mutation of the "Reader" gene ZC3H13 was higher, was detected in a total of 32 samples with mutation times of 44, accounting for 18.49% of the total m6A gene mutations. The "Writer" gene has a greater frequency of mutations than the "Reader" and "Eraser" genes, and the "Writer" gene has a higher mutation frequency as a whole (Figure 1(a)). Moreover, in the 137 READ patients, m6A regulatory gene mutations appeared in 26 independent samples (Supplementary Table 1). Among them, the mutation of the "Reader" gene ZC3H13 was higher, was detected in 7 samples, and the mutation was 12 times, accounting for 18.18% of the total m6A gene mutations. The "Writer" gene has more variability in mutation frequency than the "Reader" and "Eraser" genes, and the "Writer" gene as a whole also has a higher mutation frequency (Figure 1(b)).

However, in the 452 COAD samples with CNV data, it was observed that the m6A regulatory gene has a high frequency CNV (Supplementary Figure 1(a)). For example, the "Reader" gene YTHDF1 has the highest CNV frequency with the frequency of 76.22%, while the "Writer" gene ZC3H13 has a CNV event frequency of 69.8%, and the "Eraser" gene ALKBH5 has a frequency of 66.81% (Table 1). Meanwhile, in the 165 READ samples with CNV data, it was observed that the m6A regulatory gene has a high frequency CNV (Supplementary Figure 1(b)). For example, the "Reader" gene YTHDF1 has the highest CNV frequency, with a frequency of 87.8%. The frequency of the CNV event of the "Writer" gene ZC3H13 is 72.89%, and the "Eraser" gene ALKBH5 has a frequency of 64.63% (Table 2).

Changes in m6A regulatory genes are correlated with clinical pathology and molecular sub-characteristics

We assessed the connection between the changes in m6A modulatory genes (CNV and/or SNV) and clinicopathological features of COAD and READ people. We used univariate Cox regression analysis of each clinical feature. Results showed that the changes (CNV, CNV or SNV) of m6A regulatory gene had no significant effect on patient survival in both COAD and READ patients (Supplementary Table 2).

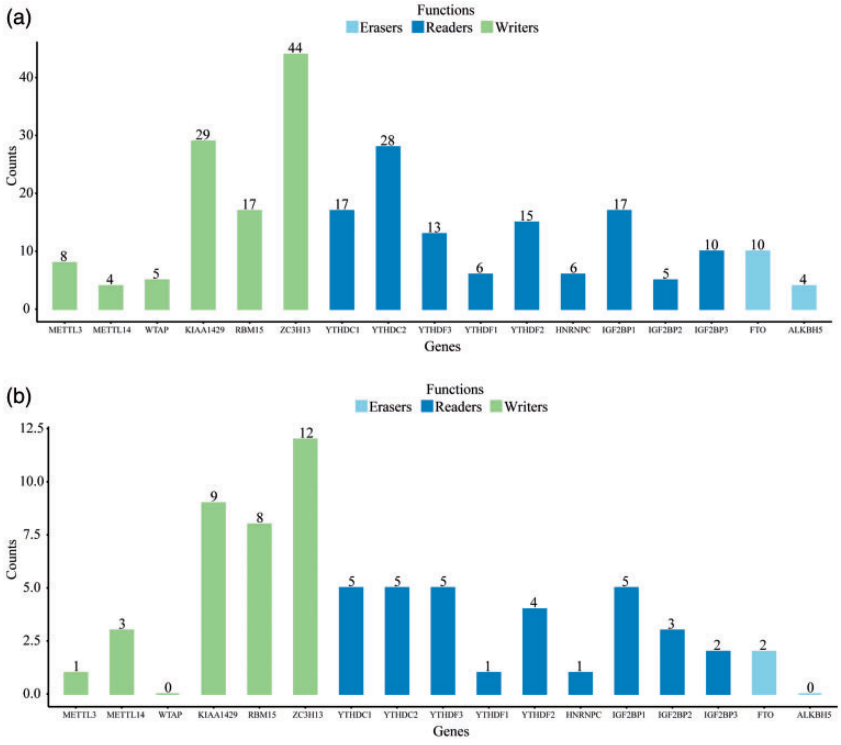

Figure 1. The mutation frequency statistics of different functional m6A regulatory genes in COAD and READ cases. (a) The distribution of SNV in COAD. (b) The distribution of SNV in READ. (A color version of this figure is available in the online journal.)

Table 1. The CNV statistics of m6A regulatory genes in COAD samples.

Type	Gene symbol	Diploid	Deletion	Amplification	CNV_sum	Amplification %	Deletion%	Percentage
Writers	METTL3	329	102	21	452	4.65	22.57	57.87
	METTL14	356	84	14	454	3.08	18.50	56.05
	WTAP	387	27	39	453	8.61	5.96	53.93
	KIAA1429	231	9	213	453	47.02	1.99	66.23
	RBM15	363	79	10	452	2.21	17.48	55.46
	ZC3H13	196	15	242	453	53.42	3.31	69.80
Readers	YTHDC1	373	81	10	464	2.16	17.46	55.44
	YTHDC2	341	97	14	452	3.10	21.46	57.00
	YTHDF3	245	10	197	452	43.58	2.21	64.85
	YTHDF1	141	1	310	452	68.58	0.22	76.22
	YTHDF2	337	114	2	453	0.44	25.17	57.34
	HNRNPC	332	102	22	456	4.82	22.37	57.87
	IGF2BP1	367	34	51	452	11.28	7.52	55.19
	IGF2BP2	389	13	55	457	12.04	2.84	54.02
	IGF2BP3	231	0	223	454	49.12	0.00	66.28
Erasers	FTO	377	12	72	461	15.62	2.60	55.01
	ALKBH5	225	223	5	453	1.10	49.23	66.81
Total		5220	1003	1500	7723	19.42	12.99	59.67

Table 2. The CNV statistics of m6A regulatory genes in READ samples.

Type	Gene symbol	Diploid	Deletion	Amplification	CNV_sum	Amplification %	Deletion%	Percentage
Writers	METTL3	101	54	8	62	12.90	87.10	38.04
	METTL14	111	50	4	54	7.41	92.59	32.73
	WTAP	124	18	22	40	55.00	45.00	24.39
	KIAA1429	66	4	96	100	96.00	4.00	60.24
	RBM15	123	36	5	41	12.20	87.80	25.00
	ZC3H13	45	6	115	121	95.04	4.96	72.89
Readers	YTHDC1	120	43	7	50	14.00	86.00	29.41
	YTHDC2	113	43	8	51	15.69	84.31	31.10
	YTHDF3	76	8	81	89	91.01	8.99	53.94
	YTHDF1	20	1	143	144	99.31	0.69	87.80
	YTHDF2	104	57	3	60	5.00	95.00	36.59
	HNRNPC	101	55	8	63	12.70	87.30	38.41
	IGF2BP1	121	19	25	44	56.82	43.18	26.67
	IGF2BP2	127	7	30	37	81.08	18.92	22.56
	IGF2BP3	69	1	94	95	98.95	1.05	57.93
Erasers	FTO	129	11	31	42	73.81	26.19	24.56
	ALKBH5	58	106	0	106	0.00	100.00	64.63
Total		1608	519	680	1199	56.71	43.29	42.71

Moreover, we observed in the previous analysis that CNV changes in the m6A regulatory gene and changes in SNV have similar characteristics in both COAD and READ patients. Although CNV or SNV was not significantly associated with patient survival outcomes, the CNV changes could affect gene expression levels through dose-compensation effects. To this end, we next evaluated the CNV effects of m6A regulatory gene on its mRNA expression. The findings revealed that mRNA expression levels were remarkably correlated with different CNV patterns in 402 COAD samples and 145 READ samples. In COAD, the increase in copy number of 14 of all 17 regulatory genes was associated with greater mRNA content, whereas deletion resulted in downregulation of mRNA content (Figure 2). In READ, an increase in the copy number of 12 genes was associated with higher mRNA expression, whereas a deletion resulted in a reduction in mRNA level (Figure 3).

For the levels of m6A regulatory genes in COAD and READ, we combined the clinical features of patients to analyze and found that only a few m6A regulatory genes have a positive correlation with the clinical pathological features (Figure 4(a) and (b), Supplementary Figures 2 to 6). Next, we focus on the correlation in the expression levels among m6A regulatory genes. The results showed that there are positive correlations between the levels of 17 m6A regulatory genes in both COAD and READ (Figure 5(a) and (b)).

Relationship between m6A-regulated genes and the patients' survival

In the previous analysis, we found a significant association between tumor stage and the life of COAD and READ people. Next, we conducted univariate Cox regression analysis on the basis of the gene expression of the m6A regulatory targets. Data showed that the RBM15, YTHDC2, and METTL14 genes in the READ sample were

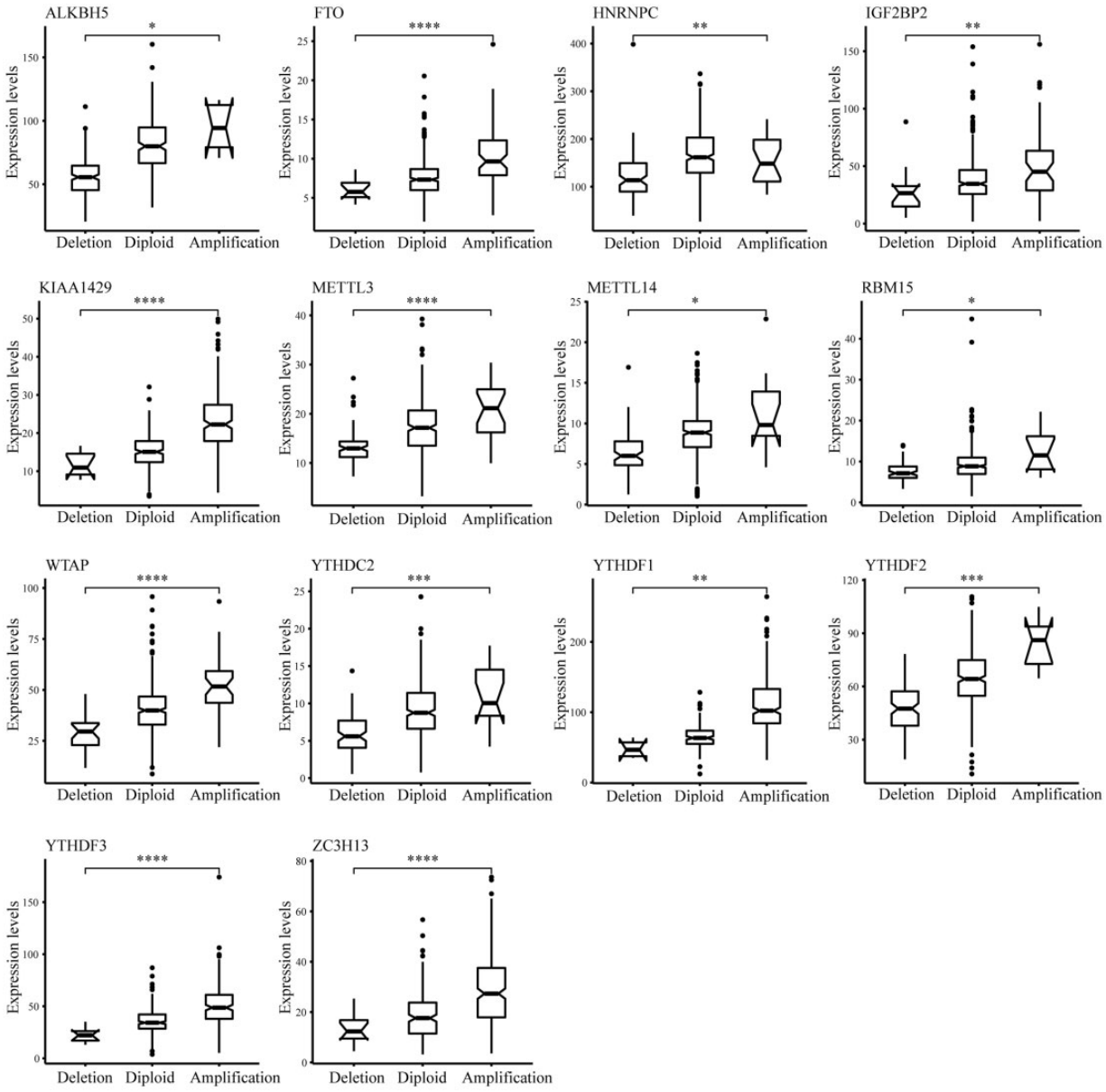

Figure 2. Connection between CNV and expression level of m6A modulatory genes in COAD. *t* test was used to examine the difference between the two groups.

vitally associated with the prognosis of patients (Supplementary Table 3).

Next, we carried out multivariate Cox analysis to calculate the risk coefficients for 17 m6A regulatory genes. The risk predictions for COAD and READ samples were based on the median risk scores. We found that, in both COAD and READ samples, the life time between high-risk and low-risk samples was dramatically different. The corresponding AUC curves have relatively high AUC values (Figure 5(c) and (d)).

Since the previous studies we have shown that CNV in the m6A regulatory gene result in changes in m6A regulatory gene expression levels. Next, we used CNV as the study object to analyze the connection between the CNV of m6A regulatory genes and patients' survival in COAD and READ. The results showed that the changes in SNV and CNV did not directly alter patients' survival in both COAD and READ samples (Supplementary Figures 7 and 8). Here, we showed the clustering of 17 m6A gene expression levels and clinical features based on risk scores obtained from multivariate Cox analysis (Figure 6(a) and (b)).

Establishment of a prognostic signature based on m6A regulatory genes

To further reduce the number of prognostic markers, we performed LASSO analysis on 17 m6A regulatory genes. Combining the results of 1000 LASSO regressions, we found that there were 2 genes that repeatedly appear in the LASSO results with 100 times in COAD samples and 80 times in READ samples. Moreover, their CNV has a significant effect on the gene expression level, resulting in YTHDC2 and IGF2BP3 (Supplementary Table 4).

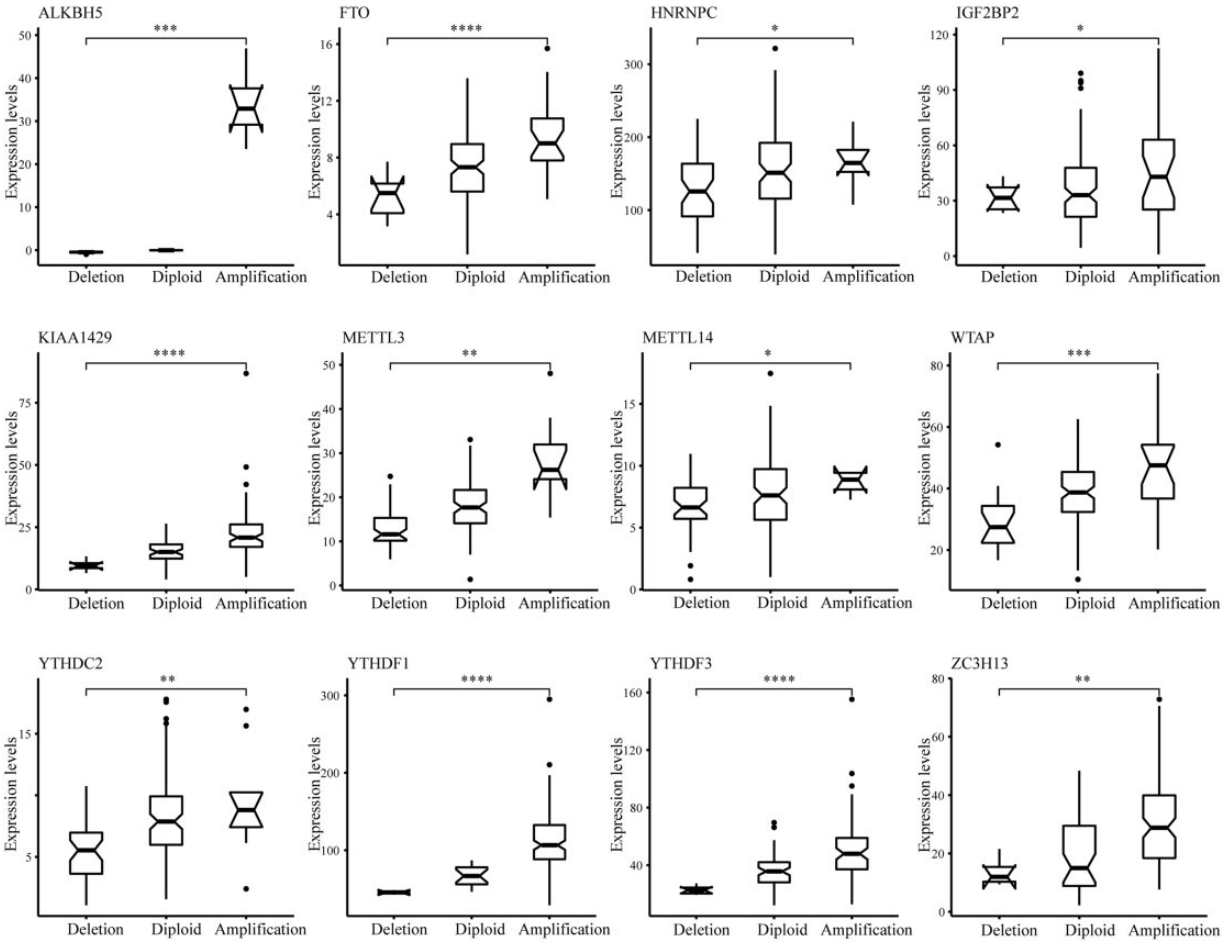

Figure 3. Connection between CNV and expression level of m6A regulatory genes in READ. *t* test was conducted to examine the difference between the two groups.

We found that these two genes were both "Reader" genes. Next, the patient's risk was predicted by the expression levels of these two genes. We used these two genes to perform multivariate Cox analysis to obtain the patient's risk scores. Using the median risk scores to predict the risk of patients, it was found that the expression of these two genes can effectively analyze and predict COAD and READ patients (Figure 6(c) and (d)). Furthermore, we selected a set of colorectal cancer data set GSE33113, containing a whole of 90 cases with follow-up information. YTHDC2 and IGF2BP3 expression profiles were extracted from the samples, and the risk score of every sample was calculated using same method and the samples were classified as high and low risk. In the external validation set, YTHDC2 and IGF2BP3 could significantly classify the patients into the high-risk population and the low-risk population, and the AUC at three years and five years was higher than 0.7 (Figure 6(e)). The one-, three-, and five-year AUCs of this signature were all greater than 0.6, and the *P*-value for risk prediction for patients was also less than 0.0001. At the same time, we performed a cluster analysis on the expression levels of these two m6A regulatory genes and their risk scores to patients, and found that different genes were preferentially expressed in patients with high-risk and low-risk (Supplementary Figure 9(a) and (b)).

Functional enrichment analysis of YTHDC2 and IGF2BP3

Given that the YTHDC2 and IGF2BP3 genes are "Reader" genes during m6A methylation, we decided to understand the roles of m6A disorders in the etiology of COAD and READ. We performed pathway enrichment analysis based on the different gene expression of YTHDC2 gene and IGF2BP3. Gene enrichment analysis implied that high IGF2BP3expression was significantly correlated with IFN-γ in COAD (Supplementary Table 5). In the READ samples, high IGF2BP3 gene expression is related to the azurophil granule membrane pathway, which is also closely related to the immune function.

Besides, gene enrichment analysis showed that high YTHDC2 gene expression in COAD is closely related to cell energy metabolism and ATP activity (Figure 7(a), Supplementary Table 6). In the READ samples, high YTHDC2 gene expression is related to the cell centrosome pathway, which is also closely related to cell division (Figure 7(b)).

Discussion

In this work, we used the gene expression, CNV, SNV and clinical messages from TCGA database. Then, 14 of all 17

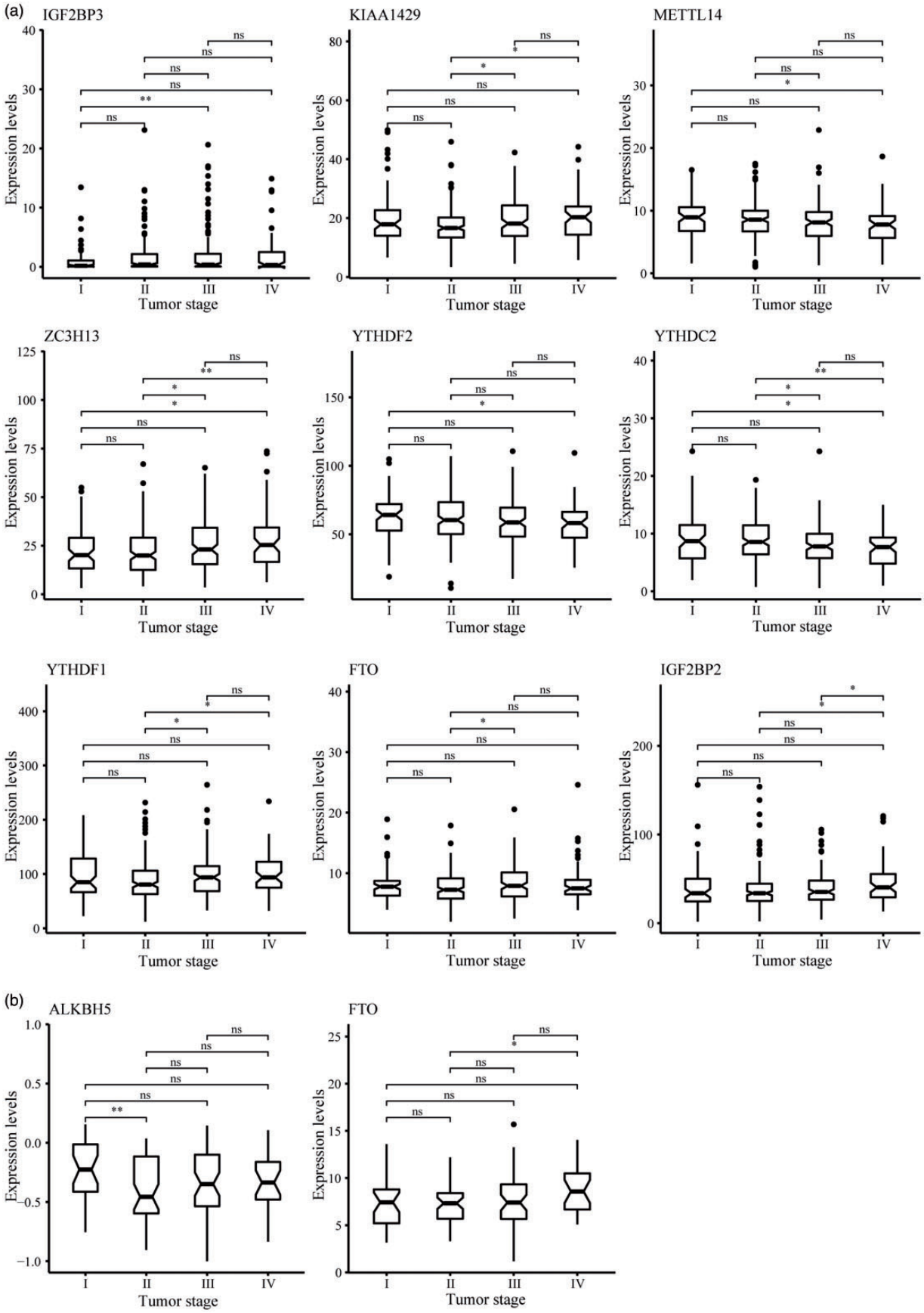

Figure 4. Relationship between gene expression level of m6A regulatory genes and tumour stage. (a) m6A regulatory genes and tumour stage in COAD. (b) m6A regulatory genes and tumour stage in READ. *t* test was used to assay the difference between the two groups.

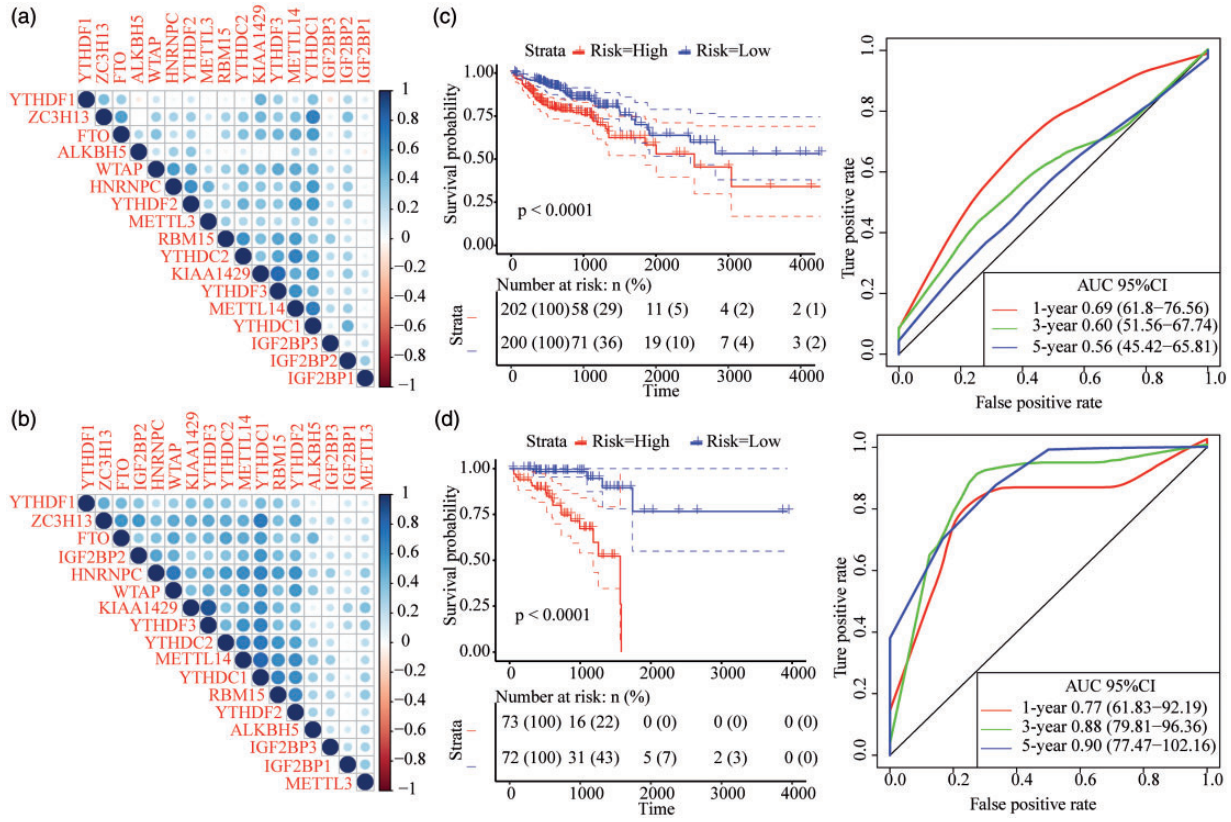

Figure 5. The m6A regulatory genes associated with clinical features and survival. (a) Correlation between levels of m6A regulatory genes in patients with COAD. (b) Correlation between levels of m6A regulatory genes in patients with READ. (c) The Kaplan–Meier curves in COAD and the ROC curves illustrated the predictive ability of the m6A regulatory genes. (d) The Kaplan–Meier curves in READ and the ROC curves illustrated the predictive ability of the m6A regulatory genes. (A color version of this figure is available in the online journal.)

m6A regulatory genes were correlated with greater mRNA content, whereas deletion resulted in decreased expression. Using univariate Cox regression analysis, RBM15, YTHDC2, and METTL14 genes in the READ samples were vitally associated with the patients' prognosis. Based on the LASSO regression models, we built a 2-gene (YTHDC2 and IGF2BP3) signature of m6A regulators with prognostic value in CRC.

Aberrant m6A RNA methylation modifications have been demonstrated to regulate carcinogenesis of various human cancers.[20] Besides, m6A methylation opens up more possibilities for early detection and intervention of cancer, which participated in biological and etiological programs, containing cellular stress, immune response, viral infection, and tissue renewal.[21] For example, the down-regulation of YTHDF1 inhibits cancer spread and susceptibility to exposure to anticancer drugs.[22] Besides, silencing YTHDF1 obviously suppressed Wnt/β-catenin pathway activity in CRC cells.[23] Li *et al.*[24] found that the knockdown of METTL3 in CRC cells greatly suppressed cell self-renewal, the frequency and stem cell migration *in vitro*, and inhibited metastasis and tumorigenesis of CRC. In another CRC study, Peng *et al.*[9] identified that upregulated METTL3 can promote metastasis via miR-1246/SPRED2/MAPK signaling pathway. Thus, these m6A regulator genes play essential function in the emergence and development of CRC.

Accumulating evidences have shown that m6A gene signatures can be prognostic or predictive factors in human cancers. A two-gene biomarker containing YTHDC2 and HNRNPC was built and could assess survival in head and neck squamous cell carcinoma people from TCGA dataset.[14] They found the levels of METTL3, YTHDF1, KIAA1429, ALKBH5, YTHDF2, METTL14, FTO, WTAP, RBM15, and HNRNPC were increased in tumor tissues, yet YTHDC2 was obviously decreased in the cancer tissues. In another study, Chai *et al.*[25] reported that seven m6A RNA methylation regulators were used in glioma to obtain risk signature. This signature was an independent prognostic indicator but could also forecast the gliomas clinicopathological characterization. However, there was no research about m6A regulator genes-related prognostic signature in CRC.

From the above analysis, we can see that COAD and READ have high similarity in the mutation and CNV of 17 m6A regulatory genes. The "Writer" gene has higher mutations than "Reader" and "Eraser". Besides, the mutation distributions of the "Writer", "Reader" and "Eraser"

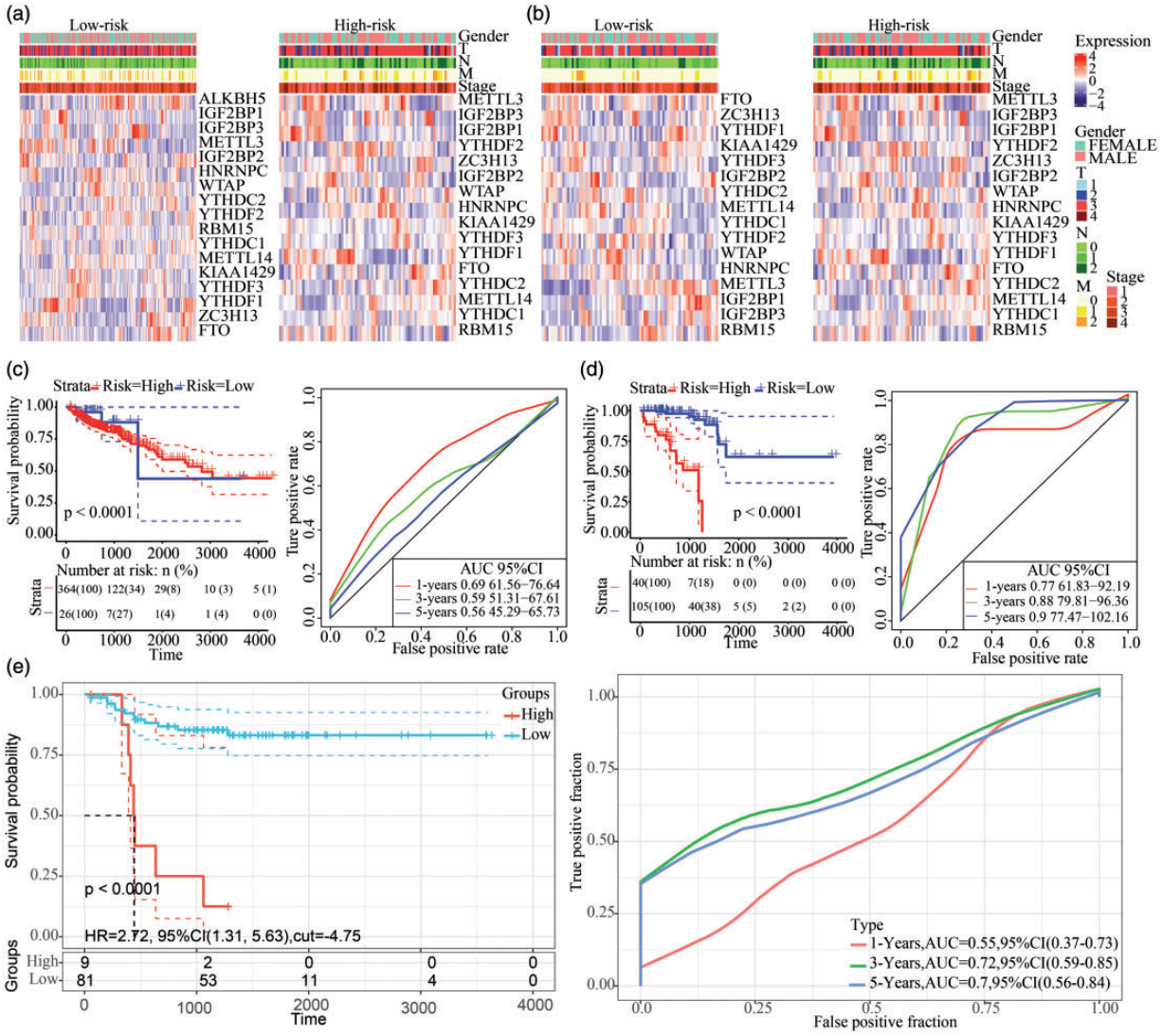

Figure 6. Establishment of a prognostic signature based on m6A regulatory genes. (a) The heat map of m6A regulatory genes and different clinical features in COAD. (b) The heat map of m6A regulatory genes and different clinical features in READ. (c) The Kaplan–Meier curves in COAD and the ROC curves assessed the predictive ability of the m6A regulatory genes by LASSO analysis. (d) The Kaplan–Meier curves in READ and the ROC curves assessed the predictive ability of the m6A regulatory genes by LASSO analysis. (e) The Kaplan–Meier curves in GSE33113 and the ROC curves assessed the predictive ability of the m6A regulatory genes by LASSO analysis. (A color version of this figure is available in the online journal.)

genes in COAD and READ are very consistent. In the CNV changes, the YTHDF1 gene in both COAD and READ samples had the highest CNV frequency, followed by the ZC3H13 and ALKBH5 genes, which were distributed among the "Writer", "Reader," and "Eraser" functions. This suggests that mutations in the m6A gene and CNV have similar changes in the development of COAD and READ.

In our signature, we identified two m6A regulator genes including YTHDC2 and IGF2BP3. YTHDC2 is known as a branch of the DExD/H-box family of ATP-dependent RNA helicases. People have determined that YTHDC2 can contribute to colon cancer metastasis by facilitating HIF-1αtranslation and as a detection marker and candidate gene.[26] Also, germline CNV in the YTHDC2 gene was reported that the YTHDC2 gene may be an underlying target for pancreatic cancer.[27] As for gene IGF2BP3, its expression has been reported to be associated with an adverse overall prognosis and metastasis in a wide variety of human carcinomas.[28] Long noncoding RNA CERS6-AS1 in combination with IGF2BP3 can improve the stability of CERS6 mRNA that plays roles in breast cancer.[29] In another study, IGF2BP3 was differentially expressed among three ocular cancers.[30] In CRC, increased IGF2BP3 expression has been shown to promote the invasive pattern of CRC in *vitro* and *vivo*.[31] Moreover, IGF2BP3 can be considered as a possible oncogene and target of miR-34a in gastric carcinogenesis.[32]

However, there are also some restrictions in the present work. First, our study is a bioinformatics and retrospective

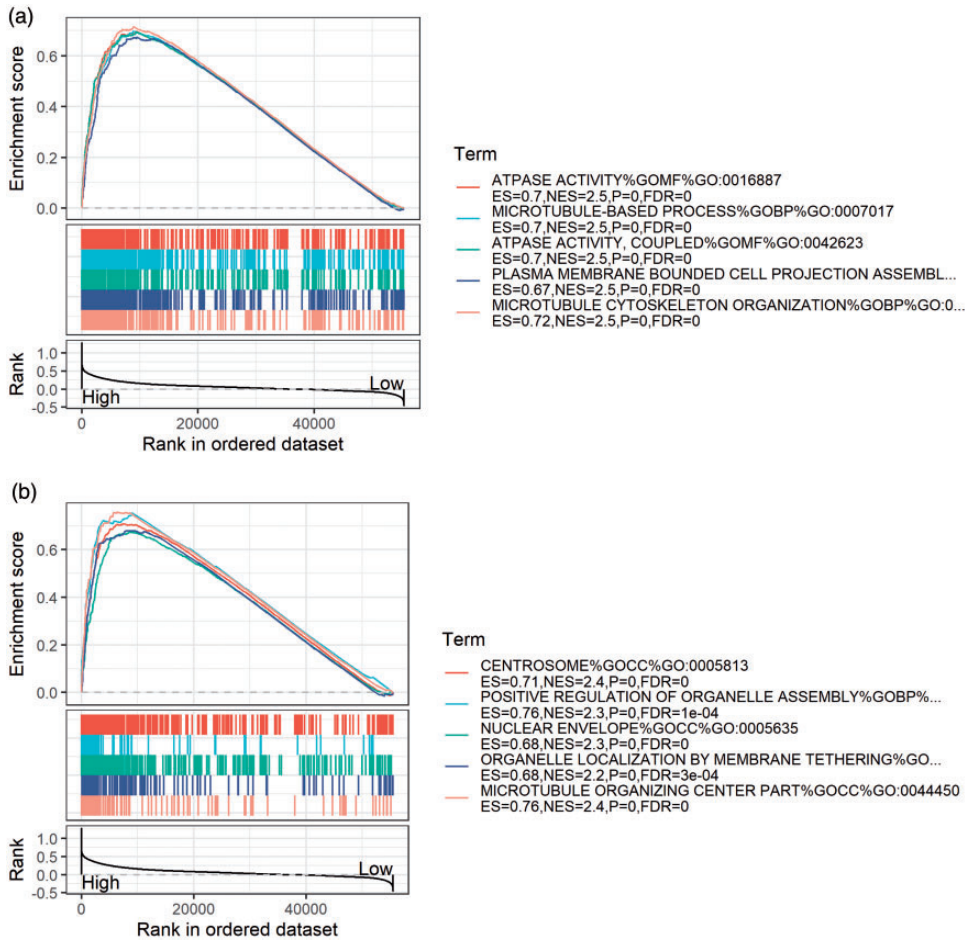

Figure 7. GSEA enrichment analysis of YTHDC2. (a) GSEA analysis in COAD. (b) GSEA analysis in READ. (A color version of this figure is available in the online journal.)

research, so the robustness of the predictive value of gene signatures should be further verified. Second, molecular experimental studies are required to validate the biological functions of gene signature in CRC.

Conclusion

In conclusion, we for the first time identified genetic alterations of m6A modulators and built a prognostic gene signature in CRC.

Authors' contributions: All authors involved in the design, interpretation of the studies and analysis of the data and review of the manuscript; JZ, XC and JW performed the bioinformatics analysis, YH analyzed the data. JY wrote the manuscript, and DG contributed to the conception of the present work and was in charge of approving the version to be published. Authors read and endorsed the final manuscript. JZ and XC contributed equally to this paper.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Dawen Guo https://orcid.org/0000-0001-5140-8004

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin* 2017;**67**:7–30
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 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
3. Hissong E, Pittman ME. Colorectal carcinoma screening: established methods and emerging technology. *Crit Rev Clin Lab Sci* 2020;**57**:22–36
4. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible m(6)a RNA methylation. *Nat Rev Genet* 2014;**15**:293–306
5. Wu F, Cheng W, Zhao F, Tang M, Diao Y, Xu R. Association of N6-methyladenosine with viruses and related diseases. *Virol J* 2019;**16**:133

6. Liu ZX, Li LM, Sun HL, Liu SM. Link between m6A modification and cancers. *Front Bioeng Biotechnol* 2018;**6**:89
7. Hong K. Emerging function of N6-methyladenosine in cancer. *Oncol Lett* 2018;**16**:5519–24
8. Hou J, Zhang H, Liu J, Zhao Z, Wang J, Lu Z, Hu B, Zhou J, Zhao Z, Feng M, Zhang H, Shen B, Huang X, Sun B, He C, Xia Q. YTHDF2 reduction fuels inflammation and vascular abnormalization in hepatocellular carcinoma. *Mol Cancer* 2019;**18**:163–63
9. Peng W, Li J, Chen R, Gu Q, Yang P, Qian W, Ji D, Wang Q, Zhang Z, Tang J, Sun Y. Upregulated METTL3 promotes metastasis of colorectal cancer via miR-1246/SPRED2/MAPK signaling pathway. *J Exp Clin Cancer Res* 2019;**38**:393–93
10. Deng R, Cheng Y, Ye S, Zhang J, Huang R, Li P, Liu H, Deng Q, Wu X, Lan P, Deng Y. m(6)a methyltransferase METTL3 suppresses colorectal cancer proliferation and migration through p38/ERK pathways. *Onco Targets Ther* 2019;**12**:4391–402
11. Ahluwalia P, Mondal AK, Bloomer C, Fulzele S, Jones K, Ananth S, Gahlay GK, Heneidi S, Rojiani AM, Kota V, Kolhe R. Identification and clinical validation of a novel 4 Gene-Signature with prognostic utility in colorectal cancer. *Int J Mol Sci* 2019;**20**:1–1
12. Sun D, Chen J, Liu L, Zhao G, Dong P, Wu B, Wang J, Dong L. Establishment of a 12-gene expression signature to predict colon cancer prognosis. *PeerJ* 2018;**6**:e4942–e42
13. Zuo S, Dai G, Ren X. Identification of a 6-gene signature predicting prognosis for colorectal cancer. *Cancer Cell Int* 2019;**19**:6–6
14. Zhao X, Cui L. Development and validation of a m(6)a RNA methylation regulators-based signature for predicting the prognosis of head and neck squamous cell carcinoma. *Am J Cancer Res* 2019;**9**:2156–69
15. Zhou J, Wang J, Hong B, Ma K, Xie H, Li L, Zhang K, Zhou B, Cai L, Gong K. Gene signatures and prognostic values of m6A regulators in clear cell renal cell carcinoma – a retrospective study using TCGA database. *Aging* 2019;**11**:1633–47
16. Cancer Genome Atlas Research N, Weinstein JN, Collisson EA, Mills GB, Shaw KR, Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. The cancer genome atlas pan-cancer analysis project. *Nat Genet* 2013;**45**:1113–20
17. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013;**31**:213–9
18. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;**33**:1–22
19. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005;**102**:15545–50
20. Sun T, Wu R, Ming L. The role of m6A RNA methylation in cancer. *Biomed Pharmacother* 2019;**112**:108613–13
21. Chen B, Li Y, Song R, Xue C, Xu F. Functions of RNA N6-methyladenosine modification in cancer progression. *Mol Biol Rep* 2019;**46**:2567–75
22. Nishizawa Y, Konno M, Asai A, Koseki J, Kawamoto K, Miyoshi N, Takahashi H, Nishida N, Haraguchi N, Sakai D, Kudo T, Hata T, Matsuda C, Mizushima T, Satoh T, Doki Y, Mori M, Ishii H. Oncogene c-Myc promotes epitranscriptome m(6)a reader YTHDF1 expression in colorectal cancer. *Oncotarget* 2018;**9**:7476–86
23. Bai Y, Yang C, Wu R, Huang L, Song S, Li W, Yan P, Lin C, Li D, Zhang Y. YTHDF1 regulates tumorigenicity and cancer stem Cell-Like activity in human colorectal carcinoma. *Front Oncol* 2019;**9**:332
24. Li T, Hu PS, Zuo Z, Lin JF, Li X, Wu QN, Chen ZH, Zeng ZL, Wang F, Zheng J, Chen D, Li B, Kang TB, Xie D, Lin D, Ju HQ, Xu RH. METTL3 facilitates tumor progression via an m(6)A-IGF2BP2-dependent mechanism in colorectal carcinoma. *Mol Cancer* 2019;**18**:112–12
25. Chai RC, Wu F, Wang QX, Zhang S, Zhang KN, Liu YQ, Zhao Z, Jiang T, Wang YZ, Kang CS. m(6)a RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. *Aging* 2019;**11**:1204–25
26. Tanabe A, Tanikawa K, Tsunetomi M, Takai K, Ikeda H, Konno J, Torigoe T, Maeda H, Kutomi G, Okita K, Mori M, Sahara H. RNA helicase YTHDC2 promotes cancer metastasis via the enhancement of the efficiency by which HIF-1alpha mRNA is translated. *Cancer Lett* 2016;**376**:34–42
27. Fanale D, Iovanna JL, Calvo EL, Berthezene P, Belleau P, Dagorn JC, Bronte G, Cicero G, Bazan V, Rolfo C, Santini D, Russo A. Germline copy number variation in the YTHDC2 gene: does it have a role in finding a novel potential molecular target involved in pancreatic adenocarcinoma susceptibility?. *Expert Opin Ther Targets* 2014;**18**:841–50
28. Lederer M, Bley N, Schleifer C, Huttelmaier S. The role of the oncofetal IGF2 mRNA-binding protein 3 (IGF2BP3) in cancer. *Semin Cancer Biol* 2014;**29**:3–12
29. Bao G, Huang J, Pan W, Li X, Zhou T. Long noncoding RNA CERS6-AS1 functions as a malignancy promoter in breast cancer by binding to IGF2BP3 to enhance the stability of CERS6 mRNA. *Cancer Med* 2019;**9**:278–89
30. Wan Q, Tang J. Exploration of potential key pathways and genes in multiple ocular cancers through bioinformatics analysis. *Graefes Arch Clin Exp Ophthalmol* 2019;**257**:2329–41
31. Xu W, Sheng Y, Guo Y, Huang Z, Huang Y, Wen D, Liu CY, Cui L, Yang Y, Du P. Increased IGF2BP3 expression promotes the aggressive phenotypes of colorectal cancer cells in vitro and vivo. *J Cell Physiol* 2019;**234**:18466–79
32. Zhou Y, Huang T, Siu HL, Wong CC, Dong Y, Wu F, Zhang B, Wu WK, Cheng AS, Yu J, To KF, Kang W. IGF2BP3 functions as a potential oncogene and is a crucial target of miR-34a in gastric carcinogenesis. *Mol Cancer* 2017;**16**:77–77

(Received March 13, 2020, Accepted May 29, 2020)