1. Introduction

Non-small cell lung cancer is an extremely common and fatal tumour, and despite some advances in the treatment of gliomas over the past decades, its outcome is still poor, and due to its strong invasive and migratory ability, it is difficult to completely resect the associated lesions by surgery, and recurrence is inevitable [1]. Therefore the development of new therapeutic approaches is of particular importance. Accurately determined molecular features of non-small cell lung cancer are needed in order to investigate further therapeutic approaches. In recent years many scholars have studied its genetics and immunometabolism [2]. Among them, the study of RNA in cancer therapy has become an important research direction, and the relationship between covalent modification of RNA and cancer has become a new research field [3].

N6-methyladenosine (m6A) modifications are associated with a variety of cancers, including solid and non-solid tumours. m6A methylation regulators play important roles in cancer development, such as proliferation, migration and invasion. Epigenetic aberrations play a crucial role in many tumours, including gliomas. n6-methyladenine (m6A) mRNA modification is the most abundant form of post-transcriptional RNA modification in eukaryotes, and a form of epigenetic modification occurs in a wide range of organisms, including viruses, yeast, plants, mice and humans. Through the rapid development of m6A RNA sequencing technology, the relevant molecular mechanisms of m6A modification have been gradually revealed. m6A methylation is currently the most common and intensively studied mode of RNA methylation [4-6]. The effects of m6A methylation modifications on RNA metabolism have been found to involve processing, nuclear export, translation, and decay [7]. Widely and dynamically changing m6A methylation. m6A methylation modifications are mainly associated with three types of proteases, which are encoded by write genes (writers), eraser genes (erasers), and read genes (readers). Among RNA modifications, N6-methyladenosine (m6A) modification is closely associated with glioma development and progression [8]. For example, the expression of m6A demethylase ALKBH5 is upregulated in non-small cell lung cancer stem-like cells, and silencing ALKBH5 inhibits GSC proliferation, thus m6A methylation plays an important role. m6A
expression is significantly reduced in non-small cell lung cancer patients, which may be related to the elevated level of FTO and the reduced level of METTL3, suggesting that the m6A regulatory factors are closely related to the occurrence of glioma [9-13]. How to effectively find molecular markers (m6A regulatory factors) to predict the prognosis of NSCLC patients is a problem we need to face.

The Cancer Genome Atlas (TCGA) is a large-scale genomic analysis of a large number of human tumours, which can be used to identify aberrations at multiple levels of genes, transcripts, proteins and epigenetics, and is increasingly being used to develop new therapeutic approaches [12]. In this study, the TCGA database was used to obtain multi-type histological data and relevant clinical features associated with non-small cell lung cancer. Bioinformatics approaches were used to screen for m6A methylation regulators that can be used as prognostic indicators in non-small cell lung cancer.

2. Materials and Methods
2.1. Data acquisition
RNA-seq transcriptome data of 1071 NSCLC patients and their corresponding clinical and prognostic information were obtained from the TCGA database. The study was conducted in accordance with the Declaration of Helsinki (revised 2013). RNA-Seq data of normal tissues from 507 normal individuals were obtained from the Genotype Tissue Expression (GTEx) (https://commonfund.nih.gov/GTEx/) database.

2.2. Expression analysis of m6A regulators
Normalised data for the number of exon model fragments per kilobase (FPKM) of RNA-Seq genes obtained from the TCGA database were converted to normalised data for the number of exon model fragments per kilobase (TPM). Differences in TPM data were then analysed using the R packages limma and edgeR. Heatmaps of the tumour patient expression data were plotted using the R package heatmap. Differentially expressed genes (DEGs) were defined with a cut-off value of \( P < 0.05 \).

2.3. Bioinformatics analysis
The Limma package (http://www.bioconductor.org/packages/release/bioc/html/limma.html) was used to analyse the relationship between the expression of m6ARNA methylation regulators and the clinicopathological variables of NSCLC, with a critical \( P \) value of 0.05. Next, the expression of 12 regulatory factors in 1071 tumour tissues and 507 normal tissues was visualised using vioplot. Spearman analysis was performed to explore the correlation between these regulatory genes. The tumour samples were then divided into two groups using the consensus clustering Plus package (https://www.bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html) and principal component analysis (PCA) was performed to validate the grouping results. Analyses of survival for both groups were handled by a survival package. To explore the prognostic role of m6A methylation modulators in NSCLC patients, we performed univariate Cox analysis and developed risk profiles by employing the least absolute shrinkage and selection operator (LASSO) Cox regression algorithm. Three genes were identified as strong prognostic factors.

2.4. Construction of prognostic risk model based on m6A methylation regulators
To determine the prognostic value of m6A methylation regulators, we performed univariate Cox regression analyses using expression data from the TCGA dataset. Least Absolute Shrinkage and Selection Operator (LASSO) regression analyses were performed to model the prognostic risk of m6A-modified genes. Coefficients were obtained by the LASSO regression algorithm. The risk score for each patient was calculated using the following formula: Risk score = (\( \text{expr}_{\text{gene1}} \times \text{coefficient}_{\text{gene1}} \) + ... +\( \text{expr}_{\text{gene}} \times \text{coefficient}_{\text{gene}} \)), where \( \text{Codfi} \) is the coefficient and \( \text{x} \) is the transformed relative expression value for each selected gene. The samples were then divided into high- and low-risk groups using the median risk score as the cutoff point. The Kaplan-Meier method was used to assess the overall survival difference between the high- and low-risk groups. The predictive effect of the model was analysed using subject work characteristics (ROC) curves and the area under the curve (AUC) was calculated.

2.5. Statistical analysis
The Wilcoxon test was used to compare the expression levels of 12m6A RNA methylation regulators between tumours and normal tissues. One-way ANOVA was used to analyse the relationship between m6A regulators and clinicopathological characteristics of HNSCC patients. The median risk score was set as a threshold value, and patients were categorised into high-risk or low-risk groups. To further analyse the OS differences between the two groups, the Kaplan-Meier method was used. R software (version 3.5.1) was used for all statistical analyses. \( P \) values less than 0.05 were considered statistically significant.

3. Results
3.1. Expression profiles of m6ARNA methylation regulators in NSCLC
To better understand the role of m6ARNA methylation regulators in carcinogenesis, we first compared the expression of m6A RNA methylation regulators between cancerous and normal tissues based on RNA data extracted from the TCGA database and identified 18 differentially expressed regulators. Both heatmap and vioplot showed RBM15B, IGF2BP2, HNRNPA2B1, IGF2BP1, YTHDF3, IGF2BP3, HNRNPC, RBM15, RBMX, METTL14, YTHDC2, METTL3, ZC3H13, WTAP, YTHDF1, and YTHDC1, YTHDF2 (\( P < 0.001 \)), and ALKBH5 (\( P < 0.01 \)) were significantly up-regulated in cancer tissues (see Figure 1A and 1B). In addition, the expression of FTO in cancer tissues was not significantly different from that in normal tissues (\( P > 0.05 \)). Analysis of these m6A-related genes showed the strongest correlations between ZC3H13 and YTHDC1 as well as HNRNPC and RBMX (see Figure 1C).

3.2. Relationship between the expression of m6A RNA methylation regulators and clinicopathological features of NSCLC patients
To investigate the relationship between m6A RNA methylation regulators and clinicopathological features of NSCLC patients, we analysed the clinical significance of these regulators individually. The results showed that the expression of RBM15B, IGF2BP2, METTL14, RBMX, HNRNPA2B1, and IGF2BP1 was significantly correle-
samples into two or more groups. We found that dividing these samples into six groups ensured significant differences between groups (see Figure 3A-D). We then used PCA to validate the classification. The results showed that clusters 1-6 were clustered separately (see Figure 3E). To further understand the relationship between clustering and clinical outcomes, we analysed the OS data for both clusters. We found that the OS of cluster 5 subgroup was higher than that of cluster 1-4, subgroup 6 (see Figure 3F).

3.4. Analysis of KEGG and GO enrichment of differentially expressed genes in different subclusters

Differentially expressed genes were screened according to the different subclusters in 2.3, and the data were analysed by functional enrichment in order to further confirm the potential functions of the potential targets. The results showed that a total of 61 genes were up-regulated and 43 genes were down-regulated in the different sub-clusters in Figure 4A, and the top 100 most differentially expressed

3.3. Cluster classification based on m6A RNA methylation expression

The Consensus Cluster Plus package was used to group 1071 NSCLC cancer tissues. Based on the cumulative distribution function values, we attempted to classify these
M6A RNA methylation in Non-Small Cell Lung Cancer

3.4. Clustering and GO analysis

Genes (40 down-regulated and 60 up-regulated) were taken to do the clustering heatmap.

A total of 2423 gene sets were found in the GO analysis, with 482 statistically significant ones, including 318 biological processes, 62 cellular components and 100 molecular functions. The biological processes were signal transduction (signal transduction), immune system process (immune system process), cell adhesion (cell adhesion). In terms of molecular function, there are protein binding, metal ion binding, and transferase activity. KEGG pathway enrichment analysis revealed that the differentially expressed genes were mainly associated with choline metabolism in cancer, EGFR tyrosine kinase inhibitor resistance, and regulation of actin cytoskeleton, and transcriptional misregulation in cancer (Figure 4B).

3.5. Prognostic role of m6A RNA methylation regulators in NSCLC

Next, we performed Cox univariate analysis to explore the prognostic role of m6A RNA methylation regulators in NSCLC. Notably, high expression of IFG2BP2, METTL14 resulted in reduced survival in NSCLC patients (hazard ratio (HR)=1.2152, 1.0637, 95% confidence interval (CI)=0.9967-1.4815, 0.8726-1.2966) (see Figure 5A). Among these m6A modifiers, IFG2BP2, and METTL14 were selected to construct risk profiles to predict prognosis based on the p-values of previous univariate analyses.

To validate the usefulness of risk profile prediction, cancer patients were categorised into high-risk and low-risk groups based on median scores. OS curves showed that the high-risk group had poorer survival than the low-risk group (see Figure 5B-D). Next, we used the coefficients of the three moderators obtained from the LASSO regression algorithm to calculate the risk scores of NSCLC patients in the TCGA dataset (see Figure 5E-G).

4. Discussion

Lung cancer is a common malignant neoplastic disease. The latest analysis of global cancer statistics shows that the incidence of lung cancer is the sixth highest in the world, with 906,000 new cases and nearly 830,000 deaths worldwide in 2020, which is the third highest mortality rate in the world [1]. Non-small cell lung cancer is the most common type of lung cancer, accounting for more than 75% of all lung cancer cases. The diagnostic and therapeutic techniques of lung cancer have improved significantly in recent years, with the diagnosis rate of early-stage lung cancer significantly increased, while the advancement of new effective targeted drugs and radiofrequency ablation techniques have effectively inhibited the progression of lung cancer and prolonged the survival of patients. However, there are fewer targeted therapeutic drugs for lung cancer and they are prone to drug resistance. The prognosis of many lung cancer patients is still unsatisfactory, and the 5-year overall survival rate is still less than 30%. Therefore, it is still of great practical significance to explore the potential mechanisms of lung cancer development and progression, to study new and effective drug targets.

Fig. 4. Differential gene expression volcano plot, differential gene KEGG pathway enrichment results and differential gene GO term enrichment results. (A) indicates the volcano plot using Fold change and corrected P-value. Red dots in the schematic indicate significantly differentially up-regulated genes, blue dots indicate significantly differentially down-regulated genes, and grey dots indicate non-significant genes. (B) is the functional enrichment results from R package ClusterProfiler(version:3.18.0), in which different colours represent the significance of the difference enrichment results, the larger the value represents the smaller the fdr value, and the size of the circle represents the number of enriched genes, and the larger the number, the larger the circle.

Fig. 5. The risk profile includes two m6A RNA methylation regulators. (A) Univariate Cox analysis of 12 m6A RNA methylation regulators in NSCLC patients. HR and 95% confidence intervals were calculated. (B-D) Kaplan-Meier OS curves of NSCLC patients assigned to high- and low-risk groups. (E-F) Risk scores of NSCLC patients in the TCGA dataset.
and therapeutic strategies, to inhibit the progression and metastasis of lung cancer, and to improve the quality of patients’ survival.

It has been shown that m6A modification plays a key role in lung cancer formation, progression and lung cancer immunity [14]. It was found that m6A methylation regulators WTAP, KIAA129 and METTL3 were significantly highly expressed in lung cancer tissues, while their value-added and invasive abilities were significantly enhanced in lung cancer cells with elevated levels of RNAm6A modification [15]. In vitro and in vivo experiments showed that the proliferation and metastasis of lung cancer were significantly elevated when METTL3 was overexpressed, while its proliferation and migration were significantly inhibited in METTL3-silenced lung cancer cells. However, some scholars found that METTL3 inhibits anti-tumour immune effects and promotes the immune escape of tumour cells by down-regulating the immune co-stimulatory molecules CD80 and CD40, which inhibit T-cell activation and aggregation in the tumour microenvironment. However, further studies have shown that inhibition of METTL3 or METTL14 improves the therapeutic efficacy of PDL inhibitors in tumour patients. Clinical studies have found that METTL14 and YTHDF1 are both significantly overexpressed in lung cancer, and their expression levels are closely correlated [16,17]. Subsequent survival analyses suggested that high expression of both METTL3 and YTHDF1 suggested a poor prognosis and lung cancer patients with high expression of both METTL3 and YTHDF1 had the worst clinical prognosis. Further gene set enrichment analyses showed that both may regulate lung cancer value-added through pathways such as RNA metabolic processes, DNA replication and cell cycle [18].

In this study, 19 M6A-related genes were extracted from TCGA. According to the expression of M6A-related genes, cluster analysis and PCA analysis were performed. The results showed that the expression levels of M6A-related genes could distinguish lung cancer patients. Cox regression analysis was used to screen two prognosis-related genes, namely IGF2BP2 and METTL14. Further LASSO algorithm determined that the key regulators of survival and prognosis of colorectal cancer patients were still IGF2BP2 and METTL14, and thus the risk prediction model was established, and colorectal cancer patients were classified into low and high-risk groups according to the median risk value. The difference in overall survival between patients in the high-risk and low-risk groups was found to be statistically significant. Univariate and multivariate Cox regression analyses showed that the risk value could be used as an independent prognostic indicator, and two prognostically relevant M6A methylation modifiers (IGF2BP2 and METTL14) were finally identified.

Next, we categorised patients into high- or low-risk groups based on their risk scores. OS curves validated that risk profiles could help differentiate patient outcomes. The results showed that the expression levels of IGF2BP2 and METTL14 tended to be higher in the high-risk group than in the low-risk group. Among them, YTHDF2 affects oncogenicity, proliferation, invasion and migration, tumour stems cell formation, metabolism and immunity, and inhibits sensitivity to chemotherapy by affecting target gene expression, and may also be involved in the regulation of non-tumour lesions such as axonal regulation, learning memory, cardiovascular disease, and toxic infections. In addition to leukaemia, YTHDF2 has been shown to promote the progression of a variety of solid tumours, including lung cancer, glioblastoma, and hepatocellular carcinoma. As a conserved RNA-binding family, the oncoproteins IGF2BP1, IGF2BP2, and IGF2BP3 are involved in a number of cellular processes, including differentiation, polarisation, proliferation, metabolism, and migration, and regulation of cancer progression. In addition, previous studies have reported the oncogenic effects of IGF2BP2 [16]. And day by day, it was observed that IGF2BP2 down-regulation has potential for tumour suppression by inhibiting tumour growth and migration. Another study also elucidated the effect of IGF2BP3 on glioma cell proliferation and invasion as well as tumour proliferation. The stable structure of METTL14 and METTL3 to form a METMERTT1 complex involved in the recognition of RNA substrates. Some studies have shown that METTL3 in the cytoplasm can also bind to ribosomes to directly promote translation of some mRNAs, such as epidermal growth factor receptor (EGFR) and some transcriptional co-stimulators of the Hippo pathway, but its function is independent of m6A methylation modification. WTAP is also an important component of the MTC complex, which localises the METTL3-METTL14 dimer to nuclear speckles and activates the function of the MTC complex to promote m6A methylation modification [19].

In summary, this study confirmed that M6A methylation regulators have significant variability in NSCLC through bioinformatics analysis, and the abnormal expression of M6A methylation regulators is closely related to the clinicopathology of NSCLC and the prognosis of patients, and the risk prediction model established in this study has important reference value for guiding the personalised treatment and improving the prognosis of NSCLC at later stages. The established risk prediction model has important reference value for guiding the personalised treatment and improving the prognosis of NSCLC, but the present study still has some limitations. For example, the sample size and clinical specimens are too small. We will expand the sample size and clinical specimens in the future, and conduct more in-depth research on the independent prognostic factors IGF2BP2 and METTL14 identified in this study.

Conflict of interests
The author has no conflicts with any step of the article preparation.

Consent for publications
The author read and approved the final manuscript for publication.

Ethics approval and consent to participate
No human or animals were used in the present research.

Informed consent
The authors declare not used any patients in this research.

Availability of data and material
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors’ contributions
Ming Zhang and Congbo Yang: Conceptualization, metho-
ology, writing original draft preparation. Wen Dong, Yu-tao Zhao and Nan Chen: Investigation, software, statistical analysis. Change Gao: Reviewing and editing, funding acquisition, supervision. All authors read and approved the final manuscript.

**Funding**

This work was supported by the National Natural Science Foundation of China [grant number 81860537], Yunnan Provincial Health Commission Medical Discipline Leader Training Project [grant number D-2019031]; Reserve Talents of Young and Middle-aged Academic and Technological Leaders Plan of Yunnan Provincial Department of Science and Technology [202105AC160026]. This work was supported by the National Natural Science Foundation of China [grant number 82160547]. Yunnan Provincial Health Commission Medical Discipline Leader Training Project [grant number H-2019073]. Yunnan Fundamental Research Projects [202101AT070143 & 202201AY070001-050].

**References**