Establishment of a 5 DNA Methylation Site Prognostic Signature Associated with N6-Methyladenosine in Lung Adenocarcinoma

Liqiang Yuan

The First Affiliated Hospital of Guangxi Medical University

Wei Jiang

People's Hospital of Guangxi Zhuang Autonomous Region

Zhanyu Xu

The First Affiliated Hospital of Guangxi Medical University

Kung Deng

The First Affiliated Hospital of Guangxi Medical University

Yu Sun

The First Affiliated Hospital of Guangxi Medical University

Junqi Qin

The First Affiliated Hospital of Guangxi Medical University

Liuliu Huang

The First Affiliated Hospital of Guangxi Medical University

Tiaozhan Zheng

The First Affiliated Hospital of Guangxi Medical University

Zhiwen Zheng

The First Affiliated Hospital of Guangxi Medical University

Fanglu Qin

Guangxi Medical University

Shikang Li (✉ shikangli@hotmail.com)

The First Affiliated Hospital of Guangxi Medical University

Research Article

Keywords: Lung Adenocarcinoma, N6-methyladenosine, DNA Methylation, Gene Signature, Prognosis

Posted Date: November 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1071199/v1>

Abstract

Background: There is a high incidence of lung adenocarcinoma (LUAD). Even with surgery, targeted therapy and immunotherapy, the survival rate of LUAD patients is still low. N6-methyladenosine (m6A) and DNA methylation markers can help with the diagnosis and treatment of LUAD patients. Therefore, it is necessary to identify a novel m6A-related DNA methylation sites signature to predict the survival of patients with LUAD.

Methods: In this study, we screened 15 m6A-related genes and their 217 methylation sites. RNA sequencing data of 15 genes and the clinicopathological parameters of TCGA-LUAD were obtained from the TCGA database (<http://cancergenome.nih.gov/>). The LUAD-DNA CpG site information was obtained from the Illumina Human Methylation 450 BeadChip (Illumina, San Diego, CA, United States). The methylation sites related to prognosis were screened using univariate COX analysis, and the independent predictors of LUAD patients were identified using multivariate COX analysis of least absolute shrinkage and selection operator (LASSO) Cox regression analysis. Finally, a model with 5 methylation sites as the main body to predict the prognosis of OS in patients with LUAD was obtained. According to the risk grouping of the prediction model, Kaplan-Meier curve and the receiver operating characteristic (ROC) curve were performed in the test and training sets to assess the predicted capacity of the model. In addition, a nomogram constructed by combining the risk score of methylation group and other related clinicopathological factors to verify the reliability of our model.

Results: We constructed a m6A-related 5-DNA methylation site model to predict OS in LUAD patients. According to the results of the Kaplan-Meier curve, both the test set and the training set, the high-risk group showed a worse prognosis. The AUCs of the 5 DNA methylation signature at 1, 5 and 10 years in test datasets were 0.730, 0.649 and 0.726, respectively, and 0.679, 0.656 and 0.732 in training datasets. Finally, we constructed a nomogram to further verify the reliability of the model.

Conclusion: In this study, we analyzed the methylation sites of m6A-related genes and established a m6A-related 5-DNA methylation site model to predict OS in LUAD patients.

1. Introduction

Based on the latest information provided by the International Agency for Research on Cancer for the year 2020, there were 19.3 million new confirmed cancer cases and 10 million cancer deaths worldwide¹. Lung cancer is the leading cause of cancer deaths. As shown in 2017, the total number of patients who died from breast, prostate and colorectal cancers combined was lower than the number of patients who died from lung cancer². Epigenetics helps put forward novel views into understand tumorigenesis³.

Epigenetics has been an attractive field in recent years⁴. Most research pays attention to DNA methylation and histone modifications⁵. However, with the development of high-throughput sequencing, RNA strands were found to contain modifications such as N6-methyladenosine (m6A), Cytosine

hydroxylation (m5C) and N1-methyladenosine (m1A) during translation^{6,7}. These modifications have proven to be closely related to RNA metabolism, especially mRNA degradation, splicing, stability and translation⁸.

Over 100 RNA modifications have been identified⁹. The m6A modification inserts a methyl substituent onto the N atom at position 6 of adenosine and is one of the most common modifications found in eukaryotic mRNA¹⁰. Among other members of the RNA family, m6A methylation can also be detected in tRNA, ribosomal RNA and microRNA¹¹. M6A methyltransferases (Writers), m6A demethylases (erasers) and m6A methylated reading proteins (Readers) regulate m6A methylation dynamically and reversibly¹².

Recently, m6A modifications have been linked to obesity, autoimmune diseases, infertility and many other diseases^{13,14}. In addition, m6A modifications are closely related to tumor progression. For example, in cervical squamous cell carcinoma (CSCC), the expression of FTO is significantly increased and the level of m6A methylation is significantly reduced, suggesting a poor prognosis¹⁵. FTO promotes the development of lung squamous cell carcinoma (LUSC)¹⁶. There is also evidence suggesting that targeting METTL3 mRNA 3'UTR binding sites inhibits the proliferation of non-small cell lung cancer (NSCLC) cells¹⁷. Therefore, a full understanding of the key m6A RNA methylation regulatory factors is critical for the treatment of lung cancer. Among all types of lung cancer, the incidence rate of lung adenocarcinoma (LUAD) is the highest¹⁸. However, little is known about the physiological process of m6A modified regulators in lung adenocarcinoma.

In our study, genomic methylation profiles of 15 widely reported m6A regulators in LUAD patients were obtained from the TCGA database, and an overall survival (OS) risk prediction model based on DNA methylation was established. Its ability as a molecular prognostic model was evaluated by ROC analysis, and a nomogram was used to determine the reliability of DNA methylation factors as LUAD biomarkers. Our conclusions will provide new methods for predicting OS in LUAD patients.

2. Materials And Methods

2.1 Select m6A modification related regulatory factors

A total of 15 genes are widely reported as m6A regulators¹⁹⁻²¹. RNAs undergo methyl group modification by "writers" such as METTL3, METTL16, WTAP, RBM15, RBM15B, KIAA1429, and ZC3H13²². Then, "readers" including YTHDF1, YTHDF3, YTHDC2, HNRNPC, HNRNPA2B1 and EIF3A recognize those m6A-modified RNAs and function in RNA processing, nuclear export, translation and decay²³. By relying on "erasers" (FTO, ALKBH5), m6A is restored to adenosine and achieves demethylation modification²⁴. RNA sequencing data of 15 genes and the clinicopathological parameters of TCGA-LUAD were obtained from the TCGA database (<http://cancergenome.nih.gov/>). The following methods were used for data screening: (1) samples without clinical data and (2) samples without survival information. Finally, a total of 455 LUAD patient samples were obtained and used for follow-up bioinformatics analysis.

2.2 DNA methylation data collection

The LUAD-DNA CpG site information was obtained from the Illumina Human Methylation 450 BeadChip (Illumina, San Diego, CA, United States) and the TCGA-biolinks package in R language was used to analyze clinical data²⁵. β values were used to visualize DNA methylation levels using the formula: $\beta = M/(U+M+100)$, where M represents methylation of the target CpG site and U represents unmethylation. A total of 485,577 methylation sites were obtained. The methylation data was matched with screened samples to evaluate the correlation between DNA methylation levels and OS.

2.3 Establishment and verification of the prediction model

The 455 samples were divided into two parts: two-thirds of the clinical samples were divided into the training group to identify and build a prognostic risk model. The remaining one-third of samples were placed into the validation group to confirm the predictive prognostic value of the biomarker. Significant methylation sites related to the prognosis of LUAD were identified using LASSO regression analysis. Meanwhile, to prevent overfitting, LASSO 1000 iterations were conducted using the “glmnet algorithm” package in R. We established a risk model based on the following formula: Risk Score = $\sum$ (Methylation level of each independent prognostic site $\times$ regression coefficient).

Patients in the training set were divided into high and low risk groups according to the risk score. Kaplan-Meier was used to draw the survival curve of high and low risk groups. A bilateral logarithmic rank test was used to test the survival difference between the high risk and low risk groups. A ROC curve was used to explore the survival of patients at 1, 5 and 10 years. In addition, the same method was used to verify the prediction model in the test set.

2.4 Establishing a predictive nomogram

According to the 'rms' package from R software, a nomogram was constructed to evaluate the reliability of the prediction tool. Other pathological parameters (gender, staging, N staging, T staging) and methylation risk score were evaluated using COX proportional risk assessment (including multivariate and univariate). Concordance index (C index), calibration chart and ROC curves were used together to evaluate the prediction ability of the nomogram.

2.5 Immune cell infiltration and biological pathways.

Immune and stromal cells are two types of cells that infiltrate the tumor immune microenvironment. We used "ESTIMATE" in the R package to evaluate the immune score, matrix score and tumor purity²⁶. Quantitative analysis of the immune cell infiltration level was observed using GSVA and Single-sample GSEA (ssGSEA) in the R language GSVA software package^{27,28}. Finally, we evaluated the 24 types of immune cells, including natural immune cells and acquired immune cells. To explore the relationship between the predictive model and immune infiltration, we analyzed the different 24 immune cells and their infiltration levels in the high risk and low risk groups ($p < 0.05$). Gene ontology (GO) and Kyoto

Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the DEGs were performed using Metascape (<http://metascape.org>).

2.5 Statistical analyses

Univariate analysis was performed in the training set to screen methylation sites significantly associated ($P \leq 0.01$) with OS in LUAD patients. Then, Lasso Cox regression analysis was used to further establish prognostic methylation sites and construct a Cox multiple regression model. A survival curve was drawn using Kaplan Meier method and was used to evaluate OS between the two groups. In this study, $p < 0.05$ was considered as statistically significant.

3. Results

3.1 Identification of the 5-methylation site signature in LUAD patients

β values of 217 methylation sites of the m6A-related genes were extracted, and 85 differential methylation sites were obtained through differential analysis. Finally, 11 methylation sites ($P \leq 0.01$) related to prognosis were obtained using univariate analysis. Then, 11 DNA methylation sites were selected as candidate prognostic factors for LUAD patients. We randomly divided the LUAD samples from the TCGA into a training dataset ($n = 320$) and a test set ($n = 135$). Next, based on Lasso Cox regression analysis and multivariate Cox proportional hazard regression, a risk score model was constructed using 11 candidate methylation sites, and finally the risk score formula of 5 methylation sites was obtained: Risk score = $10.31586047 \times \text{cg15425530} + 8.315651802 \times \text{cg01506766} - 1.566587167 \times \text{cg26651810} + 6.831866332 \times \text{cg15249062} - 2.830317036 \times \text{cg07620844}$.

With the median risk score employed as the threshold, a prognosis classifier was established to classify individuals into high- and low-risk classes in the training and test datasets (Fig. 1A–F). Results showed that the high-risk group showed worse OS than the low-risk group in the training and test datasets (Fig. 1G & H).

We measured the predictive ability of the 5 DNA methylation signatures for predicting OS by Area Under the Curve (AUC) of Time-dependent ROC curves. The AUCs of the 5 DNA methylation signature at 1, 5 and 10 years in test datasets were 0.730, 0.649 and 0.726, respectively (Fig. 1I), and 0.679, 0.656 and 0.732 in training datasets (Fig. 1J), indicating that the 5-DNA methylation signature showed strong potential to function as a prognostic marker in clinical practice.

3.2 Development and assessment of the nomogram

To further verify the ability of the 5-DNA methylation prediction model, a multivariate Cox model univariate Cox model were used for some other clinicopathological factors (gender, stage, T stage and N stage) and methylation relevant risk score (Fig. 2A & B). The results of multivariate Cox regression analysis showed that the biomarker of 5-DNA methylation significantly correlated with OS ($P < 0.05$, HR 2.58, 95% CI (1.22 ~ 5.46)) in test datasets, and the result showed in training datasets ($P < 0.05$, HR 1.59,

95% CI (1.04 ~ 2.43)), which indicated that prognostic signature is an independent prognostic predictor in LUAD patients. Then, a nomogram with risk levels and clinical risk characteristics to predict 1-, 2-, and 3-year OS incidence was constructed (Fig. 2C& D). Results showed that C index, AUC and calibration maps had higher values (Fig. 3), strongly proving the reliability of the nomogram as an important model for predicting OS of LUAD.

3.3 5-DNA methylation signature-associated immune infiltration and biological pathways

To identify the 5-DNA methylation signature-associated biological pathways, we performed differential gene expression analysis for the high- and low-risk groups, and functional enrichment analysis of DEGs was performed by metascape. A few of the top 20 pathways, including the Pro Talpha C8 complex, regulation of cell projection organization, Vitamin D receptor pathway and Chromatin modifying enzymes were observed (Fig. 4A & B). Using the GSVA package, a TCGA LUAD mRNA dataset was used to perform a single-sample Gene Set Enrichment Analysis (ssGSEA). The correlation between the different types of immune cell infiltration and 5-DNA methylation signature risk scores was analyzed using ssGSEA. Results shown in Figure 4C indicate a significant enrichment of immune cells in the low-risk group compared with the high-risk group, and the infiltration of HLA, iDCs and Type-II-IFN-Response were more obvious ($P \leq 0.01$).

4. Discussion

LUAD still contributes to high morbidity and mortality²⁹. Surgery is the best treatment of choice. However, postoperative recurrence, metastasis and even death still result in patients treated with surgery³⁰. Therefore, it is necessary to develop a new and more reliable prognostic model. A large literature has shown that multiple molecular markers can be used to determine the prognosis of a variety of cancers^{31,32}. Furthermore, DNA methylation can also be used to predict the prognosis of cancer^{33,34}. For example, Ma et al developed a prognostic model based on a 13-DNA methylation signature in patients with stage I LUAD, which can better predict OS and RFS³⁵. Similarly, DFNA5 methylation studies provide strong evidence for predicting the prognosis of breast cancer patients³⁶. In this study, we revealed the DNA methylation markers related to OS in LUAD patients based on the whole-genome methylation map of m6A-related regulatory factors.

The Kaplan-Meier method was used to analyze DNA methylation sites related to the survival and prognosis of LUAD, and the receiver operating characteristic (ROC) analysis was used to evaluate its ability as a molecular prognostic model. At the same time, we also generated a nomogram to judge the reliability of DNA methylation factors as a LUAD biomarkers. We finally developed the m6A regulatory factor-DNA methylation signature, which was able to predict the OS of LUAD patients.

In this study, final 5-methylation sites were obtained from four genes: ZC3H13, RBM15B, FTO and METTL16. ZC3H13, RBM15B and METTL16 come from the "writers" group and FTO belongs to the "erasers" group. It has been reported that these genes play an important role in the occurrence and

development of cancer. For instance, Zhu et al reported that ZC3H13 may be an upstream regulator of the Ras-ERK signal pathway, and inactivated Ras-ERK signal transduction inhibits the proliferation and invasion of colorectal cancer³⁷. Zhang et al believe that the high expression of ZC3H13 and METTL16 indicates that the survival rate of breast cancer patients is poor, while the high expression of RBM15B indicates a higher survival rate³⁸. Tao et al reported that FTO is highly expressed in bladder cancer and promotes its occurrence by regulating the MALAT / miR-384 / MAL2 axis through m6A RNA modifications³⁹. Results show that four genes related to the 5-DNA methylation signature play an important role in tumor progression.

According to the C index, ROC curves and correction maps, we successfully established a methylation nomogram of m6A-related regulatory proteins, which can effectively predict the OS of LUAD patients. Univariate COX analysis confirmed that the 5-DNA methylation signature was significantly related to the prognosis of LUAD. Combined with age, sex, tumor stage and other multivariate COX analysis, the model was more reliable to predict OS in LUAD patients. The successful establishment of the nomogram further confirmed that the prognostic model of the methylation signature of m6A-related regulatory proteins effectively predicts OS in LUAD patients. Meanwhile, the results may contribute to the development of accurate biomarkers.

Our study still showed some limitations. First, the nomogram model can also refer to other factors related to tumor prognosis in addition to the clinicopathological factors contained in the official website of the TCGA. In addition, further clinical trials are needed to verify the application value of the 5-DNA methylation signature. Moreover, the large set and complex statistical analysis of DNA methylation determination also increase the difficulty for further experimental verification. Therefore, the 5-DNA methylation signature still has a long way to go before being applied in the clinic. However, our results also have great advantages. First, we used univariate and multivariate COX analyses to screen methylation sites of the m6A-regulatory proteins related to the prognosis of LUAD, excluding other linear interferences in the study, making the results more reliable. Second, we successfully established a nomogram related to DNA methylation of the m6A-regulatory protein, and used an effective method to predict OS in LUAD patients. The exact probability of survival can be effectively judged by the nomogram, while related studies of many other molecular markers only determine the possibility of patient survival, which provides strong evidence for the application of our model in the clinic.

5. Conclusion

In summary, we obtained genome-wide methylation data for LUAD patients from the TCGA, and established an OS risk prediction model based on DNA methylation related to m6A-regulatory proteins. We hope that these results can be applied to the clinic after being confirmed by larger, more centers and more precise studies.

Abbreviations

Lung adenocarcinoma: LUAD; N6-methyladenosine:m6A; The Cancer Genome Atlas: TCGA; Overall survival: OS; Receiver operating characteristics: ROC; The area under the curve: AUC; Hazard ratio: HR. Gene Ontology: GO; Kyoto Encyclopedia of Genes and Genomes: KEGG; Single-sample GSEA: ssGSEA.

Declarations

Data Availability

RNA sequencing data of 15 genes and the clinicopathological parameters of TCGA-LUAD were obtained from the TCGA database (<http://cancergenome.nih.gov/>). The LUAD-DNA CpG site information was obtained from the Illumina Human Methylation 450 BeadChip (Illumina, San Diego, CA, United States). All data resources are open to the public. And our research adheres to the TCGA data access and publication requirements.

Acknowledgements

None.

Authors' contributions

LQY, WJ, and ZYX contributed conception and design of the study. KD and YS downloads data from the database. JQQ and LLH devised exclusion criteria for the data. LQY wrote the first draft of the manuscript. WJ and ZYX made further proofreading of the manuscript. TZZ and ZWZ consulted the relevant literature. SKL guides the above work. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding information

The study was supported by National Natural Science Foundation of China (no. NSFC81660488) and Natural Science Foundation of Guangxi Province (no. 2017GXNSFAA198123).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

There is no potential conflict of interest between the authors.

References

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209-249, doi:10.3322/caac.21660 (2021).
2. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**, 394-424, doi:10.3322/caac.21492 (2018).
3. Valencia, A. M. & Kadoch, C. Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat Cell Biol* **21**, 152-161, doi:10.1038/s41556-018-0258-1 (2019).
4. Taby, R. & Issa, J. P. Cancer epigenetics. *CA Cancer J Clin* **60**, 376-392, doi:10.3322/caac.20085 (2010).
5. Bachman, K. E. Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. *Cancer Cell* **3**, 89-95, doi:10.1016/s1535-6108(02)00234-9 (2003).
6. Motorin, Y., Lyko, F. & Helm, M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res* **38**, 1415-1430, doi:10.1093/nar/gkp1117 (2010).
7. Hussain, S. Characterizing 5-methylcytosine in the mammalian epitranscriptome. *Genome Biol* **14**, 215, doi:10.1186/gb4143 (2013).
8. Wang, H., Zhao, X. & Lu, Z. m(6)A RNA Methylation Regulators Act as Potential Prognostic Biomarkers in Lung Adenocarcinoma. *Front Genet* **12**, 622233, doi:10.3389/fgene.2021.622233 (2021).
9. Boccaletto, P. *et al.* MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res* **46**, D303-D307, doi:10.1093/nar/gkx1030 (2018).
10. R., D., K., F. & F., R. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc Natl Acad Sci U S A* **71**, 3971-3975, doi:10.1073/pnas.71.10.3971 (1974).
11. Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206, doi:10.1038/nature11112 (2012).
12. Wang, Y. *et al.* N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. *Nat Cell Biol* **16**, 191-198, doi:10.1038/ncb2902 (2014).
13. Wang, J., Chen, L. & Qiang, P. The role of IGF2BP2, an m6A reader gene, in human metabolic diseases and cancers. *Cancer Cell International* **21**, doi:10.1186/s12935-021-01799-x (2021).
14. Zhang, L. *et al.* The role of N(6)-methyladenosine (m(6)A) modification in the regulation of circRNAs. *Mol Cancer* **19**, 105, doi:10.1186/s12943-020-01224-3 (2020).
15. Zhou, S. *et al.* FTO regulates the chemo-radiotherapy resistance of cervical squamous cell carcinoma (CSCC) by targeting beta-catenin through mRNA demethylation. *Mol Carcinog* **57**, 590-597, doi:10.1002/mc.22782 (2018).
16. Liu, J. *et al.* m(6)A demethylase FTO facilitates tumor progression in lung squamous cell carcinoma by regulating MZF1 expression. *Biochem Biophys Res Commun* **502**, 456-464, doi:10.1016/j.bbrc.2018.05.175 (2018).

17. Du, M. *et al.* MiR-33a suppresses proliferation of NSCLC cells via targeting METTL3 mRNA. *Biochem Biophys Res Commun* **482**, 582-589, doi:10.1016/j.bbrc.2016.11.077 (2017).
18. Zappa, C. & Mousa, S. A. Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res* **5**, 288-300, doi:10.21037/tlcr.2016.06.07 (2016).
19. Wang, J., Zhang, C., He, W. & Gou, X. Effect of m(6)A RNA Methylation Regulators on Malignant Progression and Prognosis in Renal Clear Cell Carcinoma. *Front Oncol* **10**, 3, doi:10.3389/fonc.2020.00003 (2020).
20. Du, J. *et al.* Malignant Evaluation and Clinical Prognostic Values of m6A RNA Methylation Regulators in Glioblastoma. *Front Oncol* **10**, 208, doi:10.3389/fonc.2020.00208 (2020).
21. Zhao, Z. *et al.* The Effect of m6A Methylation Regulatory Factors on the Malignant Progression and Clinical Prognosis of Hepatocellular Carcinoma. *Front Oncol* **10**, 1435, doi:10.3389/fonc.2020.01435 (2020).
22. Tuncel, G. & Kalkan, R. Importance of m N(6)-methyladenosine (m(6)A) RNA modification in cancer. *Med Oncol* **36**, 36, doi:10.1007/s12032-019-1260-6 (2019).
23. Zhang, C., Fu, J. & Zhou, Y. A Review in Research Progress Concerning m6A Methylation and Immunoregulation. *Front Immunol* **10**, 922, doi:10.3389/fimmu.2019.00922 (2019).
24. Bayoumi, M. & Munir, M. Structural Insights Into m6A-Erasers: A Step Toward Understanding Molecule Specificity and Potential Antiviral Targeting. *Front Cell Dev Biol* **8**, 587108, doi:10.3389/fcell.2020.587108 (2020).
25. Zhou, F. *et al.* Methylation of OPCML promoter in ovarian cancer tissues predicts poor patient survival. *Clin Chem Lab Med* **52**, 735-742, doi:10.1515/cclm-2013-0736 (2014).
26. Yoshihara, K. *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* **4**, 2612, doi:10.1038/ncomms3612 (2013).
27. Hänzelmann, S. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics* **14**, 7, doi:10.1186/1471-2105-14-7 (2013).
28. Foroutan, M. *et al.* Single sample scoring of molecular phenotypes. *BMC Bioinformatics* **19**, 404, doi:10.1186/s12859-018-2435-4 (2018).
29. Travis, W. D. Pathology of lung cancer. *Clin Chest Med* **32**, 669-692, doi:10.1016/j.ccm.2011.08.005 (2011).
30. M.V., B., C.M., H. & E., O.-M. DNA methylation markers and early recurrence in stage I lung cancer. *N Engl J Med* **358**, 1118-1128, doi:10.1056/NEJMoa0706550 (2008).
31. Kashani-Sabet, M. Molecular markers in melanoma. *Br J Dermatol* **170**, 31-35, doi:10.1111/bjd.12493 (2014).
32. Dickinson, B. T., Kisiel, J., Ahlquist, D. A. & Grady, W. M. Molecular markers for colorectal cancer screening. *Gut* **64**, 1485-1494, doi:10.1136/gutjnl-2014-308075 (2015).
33. Uhl, B. *et al.* PITX2 DNA Methylation as Biomarker for Individualized Risk Assessment of Prostate Cancer in Core Biopsies. *J Mol Diagn* **19**, 107-114, doi:10.1016/j.jmoldx.2016.08.008 (2017).

34. D., D., R., L. & R., T. Analysis of DNA methylation of multiple genes in microdissected cells from formalin-fixed and paraffin-embedded tissues. *J Histochem Cytochem* **57**, 477-489, doi:10.1369/jhc.2009.953026] (2009).
35. Ma, X. *et al.* DNA methylation profiling to predict recurrence risk in stage Iota lung adenocarcinoma: Development and validation of a nomogram to clinical management. *J Cell Mol Med* **24**, 7576-7589, doi:10.1111/jcmm.15393 (2020).
36. Croes, L. *et al.* Large-scale analysis of DNMA5 methylation reveals its potential as biomarker for breast cancer. *Clin Epigenetics* **10**, 51, doi:10.1186/s13148-018-0479-y (2018).
37. Zhu, D. *et al.* ZC3H13 suppresses colorectal cancer proliferation and invasion via inactivating Ras-ERK signaling. *J Cell Physiol* **234**, 8899-8907, doi:10.1002/jcp.27551 (2019).
38. Zhang, B., Gu, Y. & Jiang, G. Expression and Prognostic Characteristics of m(6) A RNA Methylation Regulators in Breast Cancer. *Front Genet* **11**, 604597, doi:10.3389/fgene.2020.604597 (2020).
39. Tao, L. *et al.* FTO modifies the m6A level of MALAT and promotes bladder cancer progression. *Clin Transl Med* **11**, e310, doi:10.1002/ctm2.310 (2021).

Figures

Figure 1

Development of 5 DNA methylation site prognostic signature in the training set (A, C, E, G & H) and test set (B, D, F, I & J). (A & B). The heatmap displays the PSI value of a 5 DNA methylation site in signature (the color transition from green to red indicates that the PSI value increases from low to high). (C & D). The risk curve of each sample based on risk score. (E & F). The scatter plot showed the survival status of LUAD patients. (G & H). Kaplan–Meier survival curve of OS among high-risk and low-risk groups. (I & J).

Time-dependent receiver operating characteristic curves at 1, 3, 5 and 10 years based on the prognostic signature.

Figure 2

COX Multivariate Analysis and nomogram Construction of 5DNA methylation Predictive signature in training set and Test set (A & B). Univariate and multivariate Cox analysis of the 5 DNA methylation site prognostic signatures and clinical risk factors. (C & D). The nomogram for predicting proportion of patients with 1-, 2- or 3-year OS in the training set and test set.

Figure 3

The calibration plots for predicting patient 1-, 3-, 5- or 7- year OS in the training set. and test set.

Figure 4

5 DNA methylation signature immune infiltration and biological pathway enrichment in LUAD. A. Significantly enriched GO terms and KEGG pathways of parent genes. B. Network contact of GO terms and KEGG pathways. C. Difference of immune infiltration between high and low risk groups.