



## OPEN ACCESS

## EDITED BY

Rui Cao,  
Capital Medical University, China

## REVIEWED BY

Zihao Wang,  
Peking Union Medical College Hospital  
(CAMS), China  
Gang Li,  
Shandong University, China  
Pengfei Wu,  
The First Affiliated Hospital of University  
of Science and Technology of China  
Anhui Provincial Hospital, China

## \*CORRESPONDENCE

Jiheng Zhang,  
zhangjh2015@lzu.edu.cn  
Jiayang Du,  
jiayangdu@126.com  
Shaoshan Hu,  
shaoshanhu421@163.com

<sup>†</sup>These authors have contributed equally  
to this work.

## SPECIALTY SECTION

This article was submitted to RNA  
Networks and Biology,  
a section of the journal  
Frontiers in Molecular Biosciences

RECEIVED 13 May 2022

ACCEPTED 20 July 2022

PUBLISHED 26 August 2022

## CITATION

Wang N, Gao X, Ji H, Ma S, Wu J, Dong J,  
Wang F, Zhao H, Liu Z, Yan X, Li B, Du J,  
Zhang J and Hu S (2022), Machine  
learning-based screening of an  
epithelial-mesenchymal transition-  
related long non-coding RNA signature  
reveals lower-grade glioma prognosis  
and the tumor microenvironment and  
predicts antitumor therapy response.  
*Front. Mol. Biosci.* 9:942966.  
doi: 10.3389/fmolb.2022.942966

## COPYRIGHT

© 2022 Wang, Gao, Ji, Ma, Wu, Dong,  
Wang, Zhao, Liu, Yan, Li, Du, Zhang and  
Hu. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or

# Machine learning-based screening of an epithelial-mesenchymal transition-related long non-coding RNA signature reveals lower-grade glioma prognosis and the tumor microenvironment and predicts antitumor therapy response

Nan Wang<sup>1,2†</sup>, Xin Gao<sup>2†</sup>, Hang Ji<sup>3</sup>, Shuai Ma<sup>1</sup>, Jiasheng Wu<sup>2</sup>,  
Jiawei Dong<sup>1</sup>, Fang Wang<sup>1</sup>, Hongtao Zhao<sup>1</sup>, Zhihui Liu<sup>1</sup>,  
Xiuwei Yan<sup>1</sup>, Bo Li<sup>4</sup>, Jiayang Du<sup>5\*</sup>, Jiheng Zhang<sup>1,2\*</sup> and  
Shaoshan Hu<sup>1,2\*</sup>

<sup>1</sup>Department of Neurosurgery, Emergency Medicine Center, Zhejiang Provincial People's Hospital, Affiliated to Hangzhou Medical College, Hangzhou, China, <sup>2</sup>Department of Neurosurgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, China, <sup>3</sup>Department of Neurosurgery, West China Hospital, Sichuan University, Chengdu, China, <sup>4</sup>Department of Neurosurgery, Taizhou First People's Hospital, Taizhou, China, <sup>5</sup>Department of Neurosurgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China

Epithelial-mesenchymal transition (EMT) confers high invasive and migratory capacity to cancer cells, which limits the effectiveness of tumor therapy. Long non-coding RNAs (lncRNAs) can regulate the dynamic process of EMT at different levels through various complex regulatory networks. We aimed to comprehensively analyze and screen EMT-related lncRNAs to characterize lower-grade glioma (LGG) tumor biology and provide new ideas for current therapeutic approaches. We retrieved 1065 LGG samples from the Cancer Genome Atlas and Chinese Glioma Genome Atlas by machine learning algorithms, identified three hub lncRNAs including CRNDE, LINC00665, and NEAT1, and established an EMT-related lncRNA signature (EMTrLS). This novel signature had strong prognostic value and potential clinical significance. EMTrLS described LGG genomic alterations and clinical features including gene mutations, tumor mutational burden, World Health Organization (WHO) grade, IDH status, and 1p/19q status. Notably, stratified analysis revealed activation of malignancy-related and metabolic pathways in the EMTrLS-high cohort. Moreover, the population with increased EMTrLS scores had increased cells with immune killing function. However, this antitumor immune function may be suppressed by increased Tregs and macrophages. Meanwhile, the relatively high expression of immune checkpoints explained the immunosuppressive state

of patients with high EMTrLS scores. Importantly, we validated this result by quantifying the course of antitumor immunity. In particular, EMTrLS stratification enabled assessment of the responsiveness of LGG to chemotherapeutic drug efficacy and PD1 blockade. In conclusion, our findings complement the foundation of molecular studies of LGG, provide valuable insight into our understanding of EMT-related lncRNAs, and offer new strategies for LGG therapy.

#### KEYWORDS

epithelial-mesenchymal transition, long non-coding RNAs, lower-grade gliomas, tumor microenvironment, antitumor treatment

## Introduction

Lower-grade glioma (LGG), consisting of World Health Organization (WHO) grades II and III, has a better prognosis relative to glioblastoma. However, it is worth noting that LGG still has the potential for malignant transformation and strong aggressiveness (Cancer Genome Atlas Research Network, 2015; Gittleman et al., 2020). The primary treatment modality for LGG is surgical intervention with the aim of maximizing tumor resection and obtaining sufficient tissue for detailed molecular and genetic characterization (Weller et al., 2021). The WHO guidelines for the management of glioma issued in 2021 place greater emphasis on molecular diagnosis compared to the 2016 guidelines (Louis et al., 2016; Gritsch et al., 2022). This demonstrates that molecular-based research can optimize the diagnosis and treatment of LGG. Especially in the context of the breakthrough of immune checkpoint blockade (ICB) for tumor treatment, in-depth molecular studies can not only reveal the malignant mechanism of LGG but also identify patients who are sensitive to antitumor therapy.

Epithelial-mesenchymal transition (EMT), is the biological process by which epithelial cells are transformed into cells with a mesenchymal phenotype by a specific procedure. In the context of tumor formation, EMT confers a variety of characteristics associated with high malignancy to individual cancer cells (Yang et al., 2020). Current studies have demonstrated the existence of multiple complex EMT regulatory cascades in gliomas that promote malignant proliferation and metastasis of tumor cells. For example, P75CUX1 regulates EMT *via*  $\beta$ -catenin (Xu et al., 2021), and miR-19a/b promotes EMT by regulating the SEPT7-AKT-NF- $\kappa$ B pathway (Wang et al., 2021). In addition, restriction of ferritin light chain expression modulates AKT-SK3 $\beta$ - $\beta$  linked protein signaling, which in turn inhibits EMT and temozolomide resistance in glioma (Liu et al., 2020). Consequently, EMT is considered a key regulator of tumor metastatic progression and therapeutic resistance, including surgical resection, chemotherapy, radiotherapy, and targeted therapy (Dongre and Weinberg, 2019). Therefore, targeting EMT as a strategy for LGG will have broad and far-reaching effects.

Notably, lncRNAs, as one of the major regulators in EMT development, can be involved in the dynamic process of EMT at different levels through a variety of complex regulatory networks (Liu et al., 2021). In hepatocellular carcinoma, the amplification of

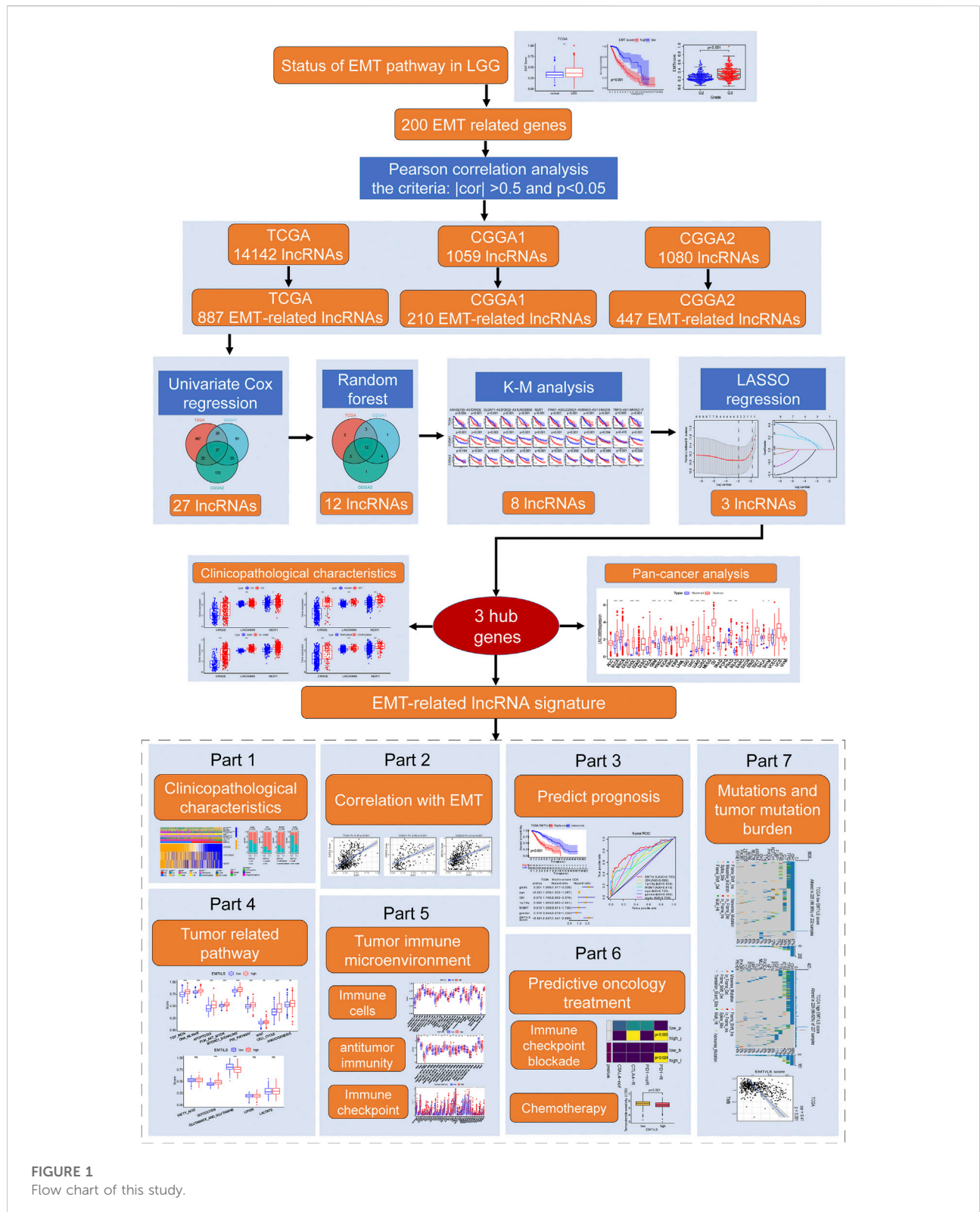
lncRNA ZFAS1 promoted the expression of EMT-related genes (Li et al., 2015). STAT3-activated lncRNA HOXD-AS1 suppressed the expression of SOX4 by blocking miR-130a-3p, which eventually upregulated EMT-related signaling targets and similarly enhanced the migration and invasion of hepatocellular carcinoma (Wang et al., 2017). Since most studies to date have focused only on the specific functions of lncRNAs, our overall understanding of EMT-related lncRNAs is limited, especially concerning the impact on LGG. Therefore, a comprehensive analysis of EMT-related lncRNAs is urgently needed. Furthermore, the understanding of how EMT is regulated to affect the tumor microenvironment (TME) is steadily increasing (Dongre and Weinberg, 2019). A study has shown that EMT-related gene expression may alter the level of T-cell infiltration, which may affect the responsiveness of immunotherapy and patient survival (Wang et al., 2018). Therefore, an in-depth analysis of the TME of EMT-related lncRNAs in LGG will provide new strategies for tumor biology and tumor therapy research.

Accordingly, we determined that EMT is significantly activated and affects the prognosis of LGG. Utilizing machine learning, three hub EMT-related lncRNAs were deeply mined and used as markers to construct an EMT-related lncRNA signature (EMTrLS) to further investigate the different features of mRNA expression profiles, clinicopathological parameters, malignant pathways, tumor metabolism, gene mutations, and tumor mutation burden (TMB) in LGG. In addition, stratified analysis of the EMTrLS scores quantified LGG immune cell infiltration and antitumor immune function. Notably, in addition to its strong prognostic value, the EMTrLS score is also intensely sensitive in predicting the efficacy of chemotherapy and treatment response to ICB in LGG patients. In conclusion, our analysis of EMT-related lncRNAs quantified the characteristics of LGG and provided a viable reference for the targeted treatment of LGG.

## Materials and methods

### Gene expression dataset

The LGG transcriptome expression profiles [Fragments Per Kilobase of exon model per Million mapped reads (FPKM) values] and corresponding clinicopathological data of The Cancer Genome



**FIGURE 1**  
Flow chart of this study.

Atlas (TCGA) dataset were obtained from the TCGA GDC project of the UCSC Xena data portal (<https://xena.ucsc.edu/>). After removing samples with no survival information or a survival

time of fewer than 30 days, 475 LGG samples were finally obtained. mRNA expression profiles (FPKM) of 105 normal brain tissues from GTEx used as a control with TCGA LGG

were also obtained from UCSC Xena. To make the gene expression profiles comparable across platforms, we converted the two sets of FPKM values to Transcripts Per Kilobase of exon model per Million mapped reads (TPM) values using the R package “limma” (Ritchie et al., 2015). The processed values were then fused to reduce batch effects and errors introduced by the integration process. The transcriptome expression profiles of 33 cancers were obtained also from the TCGA GDC project of the UCSC Xena data portal.

In addition, somatic mutation data (MAF format) were downloaded from the TCGA database (<https://portal.gdc.cancer.gov/>) for 475 LGG patients. The mutation types and frequencies of genes were analyzed and visualized using the “maftools” R software package (Mayakonda et al., 2018). The somatic mutation data were also used to calculate TMB, which can predict immunotherapy response (Chalmers et al., 2017).

The CGGA1 dataset was obtained from the CGGA 325 project in the China Glioma Genome Atlas (CGGA) database (<http://www.cgga.org.cn/>). The CGGA2 dataset was obtained from the CGGA 693 project. Based on the same inclusion criteria, the expression profiles and corresponding clinical data of 170 and 420 LGG samples were identified in CGGA1 and CGGA2, respectively (Supplementary Table S1). In addition, 20 CGGA nonglioma samples (Dataset ID: mRNA sequencing control) were included as controls. The data was also normalized by TPM and batch effects were removed with the aid of the R package “limma”.

## Analysis flow chart

After data acquisition and processing, we followed the flow chart in Figure 1 for further analysis.

## Epithelial-mesenchymal transition-related lncRNAs

The TCGA, CGGA1, and CGGA2 datasets were annotated based on the Genome Reference Consortium Human Build 38 (GRCh38) annotation file from the GENCODE website (<https://www.genecodegenes.org/human/>). A total of 14,142, 1,059, and 1,080 lncRNAs were identified in the three datasets, respectively. We retrieved a total of 200 EMT signature genes from the Molecular Signatures Database (MSigDB, <https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>). Based on the criteria of an absolute value of Pearson correlation coefficient >0.5 and *p*-value < 0.05, we screened 887, 210, and 447 EMT-related lncRNAs in TCGA, CGGA1, and CGGA2, respectively.

## Constructing EMTrLS

Combined with survival information, all EMT-related lncRNAs were screened by univariate Cox regression

analysis, and 539, 133, and 210 lncRNAs with prognostic values were obtained in the TCGA, CGGA1, and CGGA2 datasets, respectively. Twenty-seven of these lncRNAs had significant prognostic significance in all three datasets. The random forest (RF) algorithm was used to rank the weights of the prognostic value of the 27 lncRNAs with the R package “randomForest”. Only 12 of the same gene importance values in the three datasets ranked among to the top 20, and Kaplan-Meier (K-M) survival curves showed that 8 of the 12 had a significant prognostic impact. Finally, least absolute shrinkage and selection operator (LASSO) regression analysis of the eight lncRNAs in the three datasets was performed by the R package “glmnet” to screen the biomarkers for EMTrLS. Based on the LASSO regression coefficients and expression [after log<sub>2</sub> (TPM+1) transformation] of each hub gene, the following equation was used to calculate the EMTrLS score for LGG:

$$EMTrLS\ score = \sum_{i=0}^n (Regression\ Coefficient * Expression)$$

For the stratified analysis, samples from the three datasets were grouped into EMTrLS-high and -low groups according to the median value of TCGA.

Receiver operating characteristic (ROC) curves and corresponding area under the curve (AUC) values were used to assess the predictive power of the EMTrLS score and clinicopathological characteristics on prognosis using the “survivalROC” R package. The “survival” package was used to perform univariate and multivariate Cox regression analyses to assess the independent prognostic value of EMTrLS and clinicopathological characteristics. The results were visualized with the “forestplot” package.

## Functional enrichment analysis

The R package “limma” was used to screen for differentially expressed genes (DEGs) between the EMTrLS-low and -high groups, with the implementation criteria of |log<sub>2</sub>(Fold Change)| > 1, *p* < 0.05, and FDR < 0.05. Based on the R packages “clusterProfiler” and “enrichplot”, the 913 DEGs highly expressed in the high EMTrLS score group [log<sub>2</sub>(Fold Change) > 1] underwent for gene ontology (GO) [cellular component (CC), molecular function (MF), biological process (BP)] and Kyoto Encyclopaedia of Genes and Genomes (KEGG) analyses.

## Single-sample gene set enrichment analysis

The single-sample gene set enrichment analysis (ssGSEA) algorithm is calculated by rank normalization of the sample



gene expression values based on genes in a known pathway and then utilization of an empirical cumulative distribution function to determine the enrichment status in that pathway (Barbie et al., 2009). With this calculation, we explored the activation status of nine classical tumor pathways and five metabolism-related pathways. The marker genes for the pathways of interest were obtained from MSigDB. ssGSEA scores were normalized to a unit distribution in the analysis, where 0 is the minimum activation value for each pathway and 1 represents the maximum.

## Immunoinformatics analysis

The “estimate” R package was used to assess the overall level of immune infiltration for each sample, ultimately providing three scores: immune score, stromal score, and estimated score (Yoshihara et al., 2013). We used two methods to assess the abundance of each type of immune cell infiltration within the tumor tissue, namely, TISIDB (<http://cis.hku.hk/TISIDB>) and ImmuCellAI (<http://bioinfo.life.hust.edu.cn/ImmuCellAI>). TISIDB is a user-friendly website containing signature genes for 28 tumor-infiltrating immune cells (Ru et al., 2019). With these signature genes, we quantified the immune infiltration of LGG samples using the ssGSEA algorithm. The ImmuCellAI website enables online analysis and estimation of immune infiltration for each sample, including the total infiltration fraction and the level of infiltration of 24 immune cell types (Miao et al., 2020).

Tumor Immunophenotyping (TIP) (<http://biocc.hrbmu.edu.cn/TIP/>) is a tool that allows easy and rapid analysis of the anticancer immune process (Xu et al., 2018). It divides this process into 7 steps: step 1 tumor cell antigen release, step 2 cancer antigen presentation, step 3 stimulation and activation, step 4 immune cell transfer to the tumor, step 5 immune cell infiltration, step 6 T-cell recognition of cancer cells and step 7 killing of cancer cells. By online analysis, we obtained quantitative results of the anticancer immune process for each LGG sample.

## Predicting the responsiveness of anticancer treatments

The submap algorithm in GenePattern (<https://cloud.genepattern.org/gp/pages/index.jsf>) was used to predict the response of EMTrLS-high and -low cohorts to ICB treatment. The advantage of this algorithm is that it uses an unsupervised bidirectional subset projection approach that reveals similar subtypes between independent datasets (Hoshida et al., 2007). Gene expression files and information files for the immunotherapy samples used as controls were obtained from a report on melanoma (Roh et al., 2017). After obtaining the

normal  $p$ -values and Bonferroni-corrected  $p$ -values, they were visualized using the R package “pheatmap”.

The R package “PRRophetic” contains the drug effects of various cell lines included in the Cancer Genome Project (CGP) database. Based on this, a ridge regression analysis was constructed to predict the IC50 value of the drug by combining LGG expression profiles and high and low EMTrLS cohort grouping information; the smaller the IC50 value was, the stronger the drug’s ability to inhibit tumor cell growth.

## Statistical analysis

Statistical analyses were performed based on using R software version 4.1.2. K-M analysis and log-rank statistical tests were used to detect differences in overall survival (OS) between the groups. The Wilcoxon test was used to compare differences in EMT, EMTrLS scores, ssGSEA scores, and gene expression. The Kruskal test was used to compare differences between histopathological subtypes. Fisher’s exact test was used to assess differences in O (6)-methylguanine DNA methyltransferase (MGMT) promoter methylation status, IDH status, WHO grade, and 1p/19q status between the low and high EMTrLS score subgroups in response to ICB treatment.  $p < 0.05$  was considered to be statistically significant.

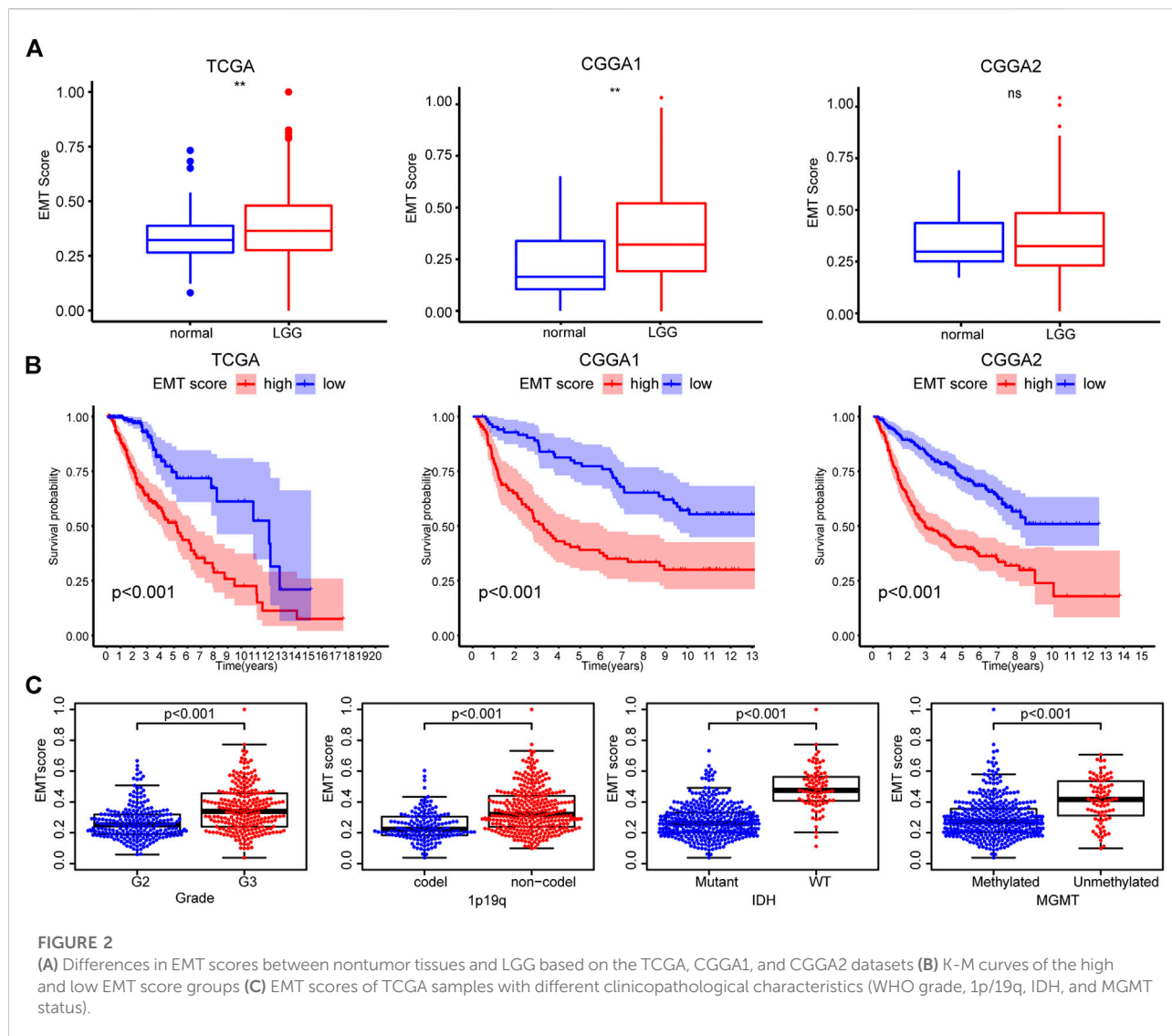
## Results

### Characteristics of Epithelial-mesenchymal transition

First, we evaluated the activation of the EMT pathway in LGG using the ssGSEA algorithm. A higher EMT score was found in LGG tissue than in normal tissue, which was validated in the CGGA1 dataset (Figure 2A). Interestingly, LGG patients with increased EMT scores had decreased OS compared with patients with low scores (Figure 2B). Moreover, samples characterized by WHO grade II, IDH mutation, and 1p19q codeletion had a significantly decreased EMT score, corroborating the association between the EMT score and poor prognosis (Figure 2C, Supplementary Figures S1A,B). The above results suggested an aberrant EMT in LGG tumor cells and an association between increased activation of EMT and enhanced malignancy of LGG.

### Machine learning screening of Epithelial-mesenchymal transition-related lncRNAs

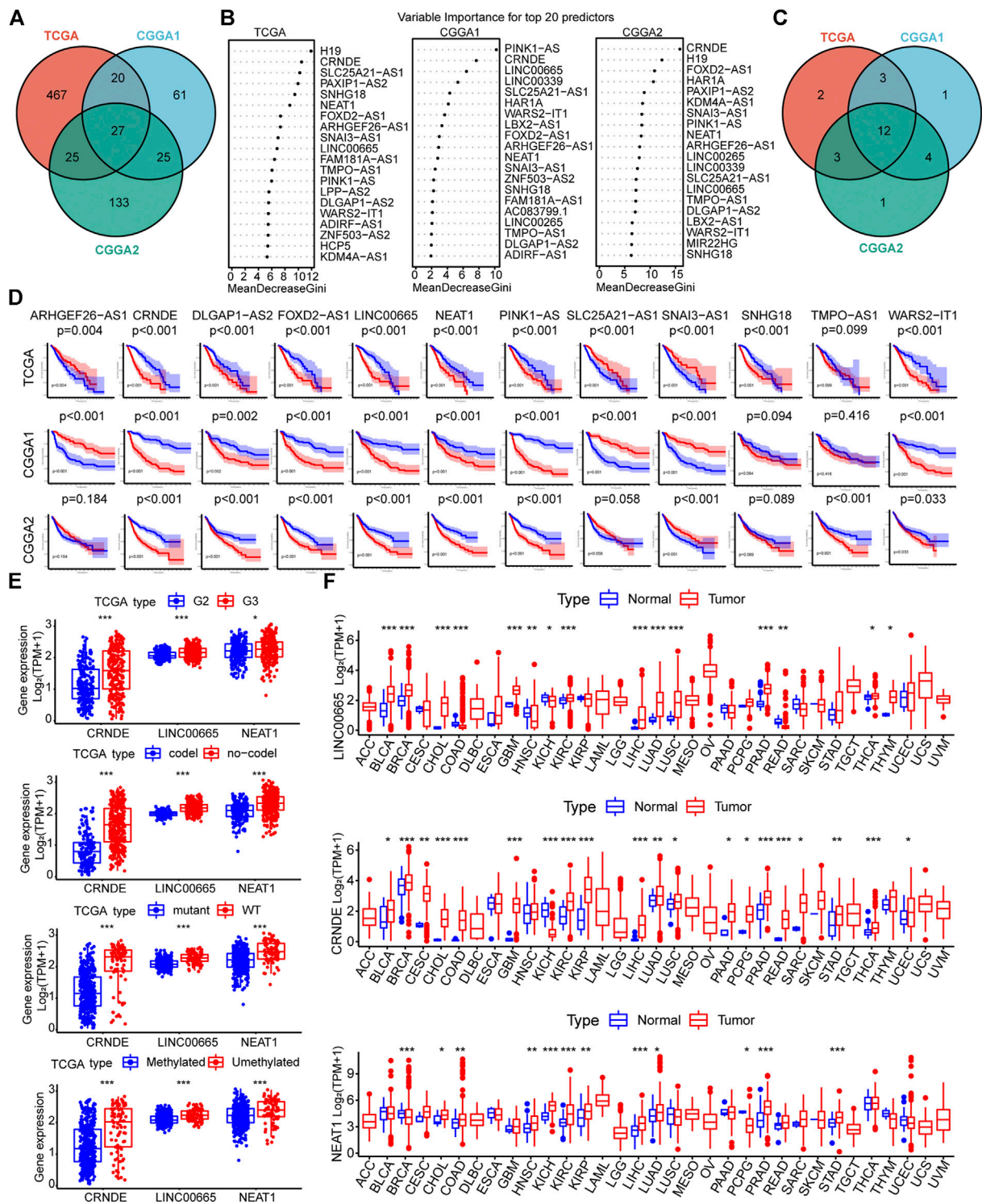
Then, lncRNAs associated with the 200 EMT signature genes were identified using Pearson correlation analysis. As a result, we



obtained 887, 210, and 447 EMT-related lncRNAs ( $| \text{cor} | > 0.5$ ,  $p < 0.05$ ) in the TCGA, CGGA1, and CGGA2 datasets, respectively (Supplementary Tables S2–S4). Furthermore, 539, 133, and 210 lncRNAs of prognostic significance were identified, and 27 of them were common among the three datasets (Figure 3A, Supplementary Table S5). The RF algorithm allows for feature importance assessment, and we focused on assessing how much each lncRNA contributes to the prognosis and comparing the importance of different EMT-related lncRNAs after determining the Gini mean. Notably, only 12 lncRNAs ranked in the top 20 in the three datasets based on the mean Gini decrease (Figures 3B,C). The K-M curves of 12 lncRNAs showed that the high and low expression group of CRNDE, DLGAP1-AS2, FOXD2-AS1, LINC00665, NEAT1, PINK1-AS, SNAI3-AS1, and WARS2-IT1 showed significant differences in survival time (Figure 3D). Accordingly, we

performed LASSO regression to screen for significantly characterized genes again and finally identified three hub lncRNAs, namely CRNDE, LINC00665, and NEAT1 (Supplementary Figure S1C).

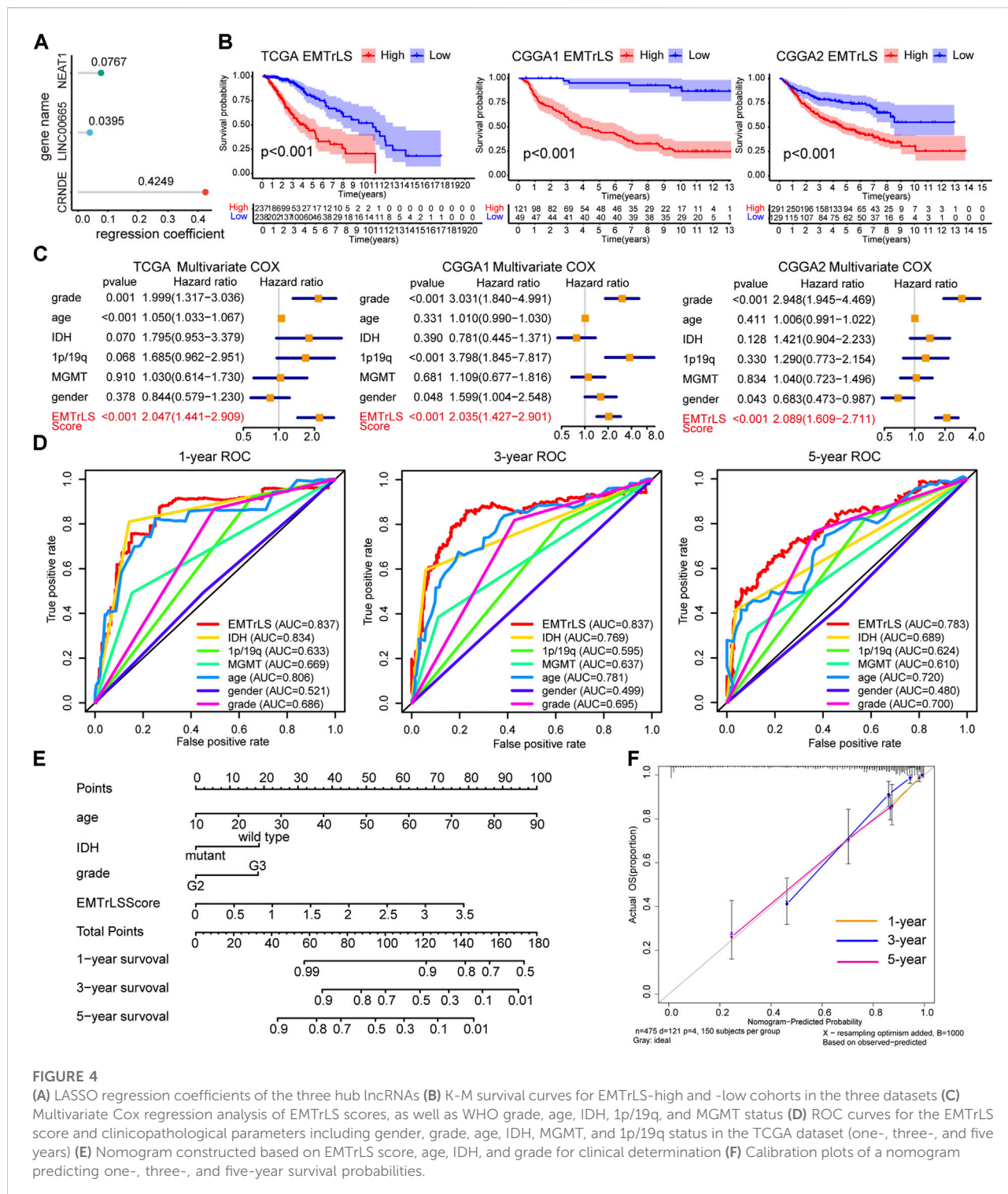
Furthermore, we explored the association between the expression of the three hub EMT-related lncRNAs and different clinicopathological features. The expression of the hub lncRNAs was down-regulated in samples with IDH mutation and 1p19q codeletion. CRNDE showed low expression in WHO grade II samples, while the expression of LINC00665 and NEAT1 remained stable across glioma grades in the CGGA cohorts (Figure 3E, Supplementary Figures S1D,E). In addition, the expression of the hub lncRNAs did not seem to be strongly correlated with MGMT promoter methylation status. Next, pan-cancer analysis was employed to investigate the effect of the hub



**FIGURE 3**

(A) Venn diagram showing the results of univariate Cox regression analysis of EMT-related lncRNAs in the three datasets (B) Importance ranking of EMT-related lncRNAs based on Gini coefficient of the random forest algorithm (top 20) (C) Venn diagrams of the top 20 lncRNAs in the three datasets in terms of prognostic impact (D) K-M curves of 12 lncRNAs. The samples were divided into two groups based on the median value of each gene expression. Red represents the high expression group, and blue represents the low expression group (E) Expression of the three hub lncRNAs in TCGA samples with different clinicopathological parameters (WHO grade, 1p/19q, IDH, and MGMT status) (F) Pan-cancer analysis of the three hub EMT-related lncRNAs (\**p* < 0.05, \*\**p* < 0.01, and \*\*\**p* < 0.001).





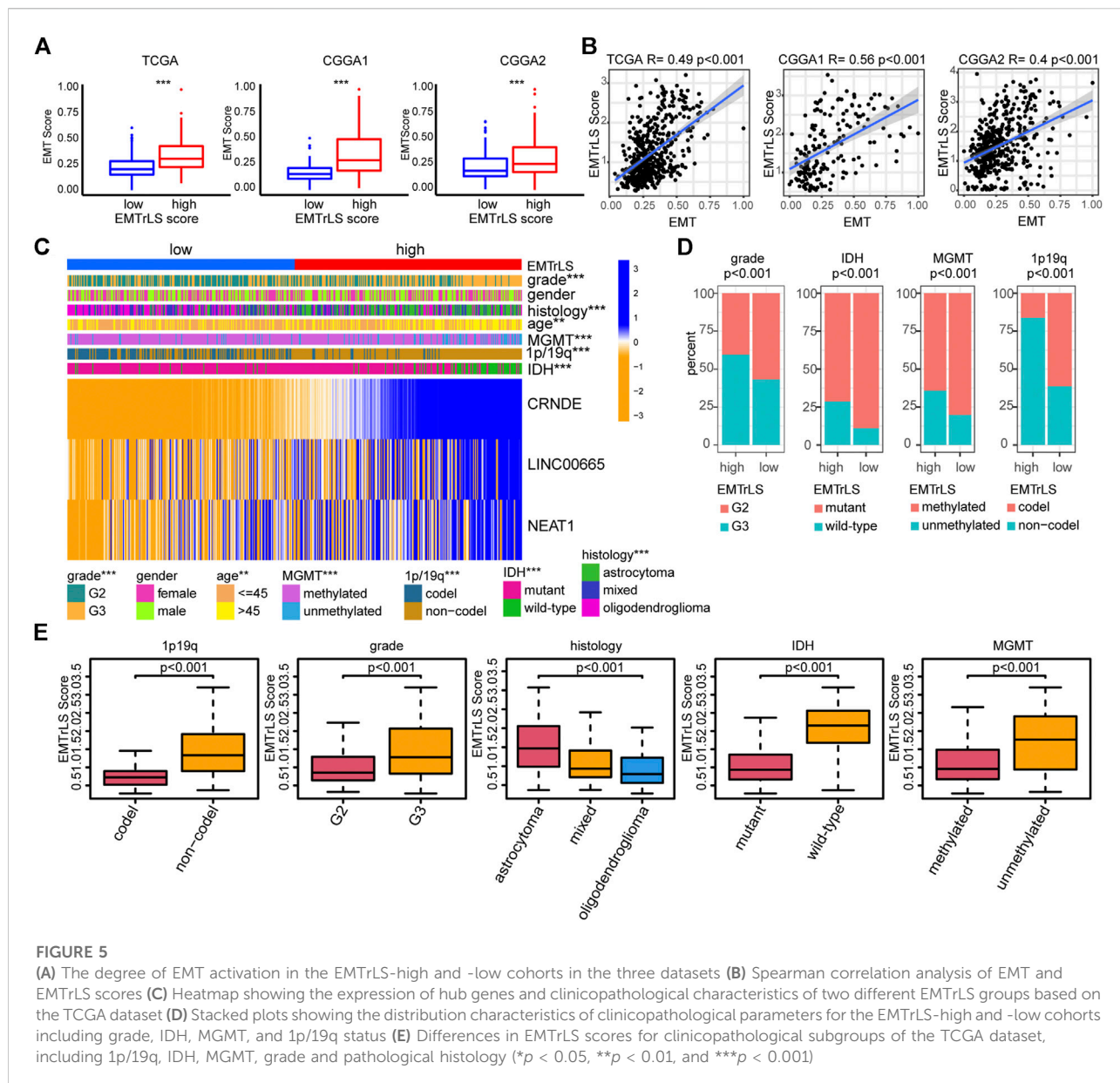
**FIGURE 4**

(A) LASSO regression coefficients of the three hub lncRNAs (B) K-M survival curves for EMTrLS-high and -low cohorts in the three datasets (C) Multivariate Cox regression analysis of EMTrLS scores, as well as WHO grade, age, IDH, 1p/19q, and MGMT status (D) ROC curves for the EMTrLS score and clinicopathological parameters including gender, grade, age, IDH, MGMT, and 1p/19q status in the TCGA dataset (one-, three-, and five years) (E) Nomogram constructed based on EMTrLS score, age, IDH, and grade for clinical determination (F) Calibration plots of a nomogram predicting one-, three-, and five-year survival probabilities.

lncRNAs on tumorigenesis development. The results showed that LINC00665, NEAT1, and CRNDE were differentially expressed in a variety of tumors compared with corresponding normal tissues, such as BRCA, CHOL,

COAD, KICH, KIRC, KIRP, LICH, LUAD, and PRAD (Figure 3F). Taken together, these results suggested that the three hub lncRNAs associated with EMT are potential indicators for quantifying the malignant features of tumors.





## Prognostic value of Epithelial-mesenchymal transition-related LncRNA-based stratification

Samples' EMTrLS was quantified based on the expression of the three hub lncRNAs and their LASSO regression coefficients (Figure 4A). K-M survival curves showed that samples with increased EMTrLS scores had significantly decreased OS ( $p < 0.001$ ) (Figure 4B). Univariate and multivariate Cox regression analyses suggested that EMTrLS could be an independent risk factor (Figure 4C, Supplementary Figure S2A). In addition, ROC curves were used to assess the

sensitivity of prognostic indicators, showing that the one-, three-, and five-year AUC values for the EMTrLS score were greater than those of molecular markers such as IDH, 1p/19q, and MGMT, respectively. This result suggested a robust time-dependent predictive power of the EMTrLS (Figure 4D, Supplementary S2B,C). In view of the robust prognostic ability of EMTrLS, we constructed clinical nomogram plots to enable optimization of clinical treatment (Figure 4E). Meanwhile, the calibration plots validated the predictive power of the nomogram with powerful accuracy (Figure 4F). In conclusion, the EMTrLS constructed using the three hub lncRNAs was of prognostic significance.

## Correlation between EMTrLS and Epithelial-mesenchymal transition scores

To further validate the correlation between our constructed EMTrLS and tumor EMT status, we first explored the difference in EMT scores between the two LGG groups with high and low EMTrLS scores. The results showed that EMT scores were significantly increased in the EMTrLS-high samples, implying activation of EMT ( $p < 0.001$ ) (Figure 5A). Furthermore, we performed a Spearman correlation analysis between EMTrLS and EMT scores, and the correlation coefficients were: 0.49, 0.56, and 0.4 in the TCGA, CGGA1, and CGGA2 datasets, respectively ( $p < 0.001$ ) (Figure 5B). In summary, EMTrLS is associated with EMT activity.

## Association between EMTrLS and clinicopathological features

With the heatmap, we further demonstrated the expression of the three hub genes in EMTrLS and found that the expression of the three genes increased with the EMTrLS score (Figure 5C, Supplementary S3A,B). Furthermore, we found that patients with low EMTrLS scores were characterized by IDH mutation, WHO grade II, 1p/19q codeletions, and MGMT promoter methylation ( $p < 0.001$ ) (Figure 5D). In addition, the EMTrLS scores also showed significant differences when IDH, WHO grade and 1p/19q status differed, which was perfectly validated by the two CGGA datasets. Notably, the EMTrLS scores also differed significantly by LGG pathology histology ( $p < 0.001$ ) (Figure 5E, Supplementary S3C,D). In summary, EMTrLS can distinguish LGGs with different characteristics.

## Functional enrichment

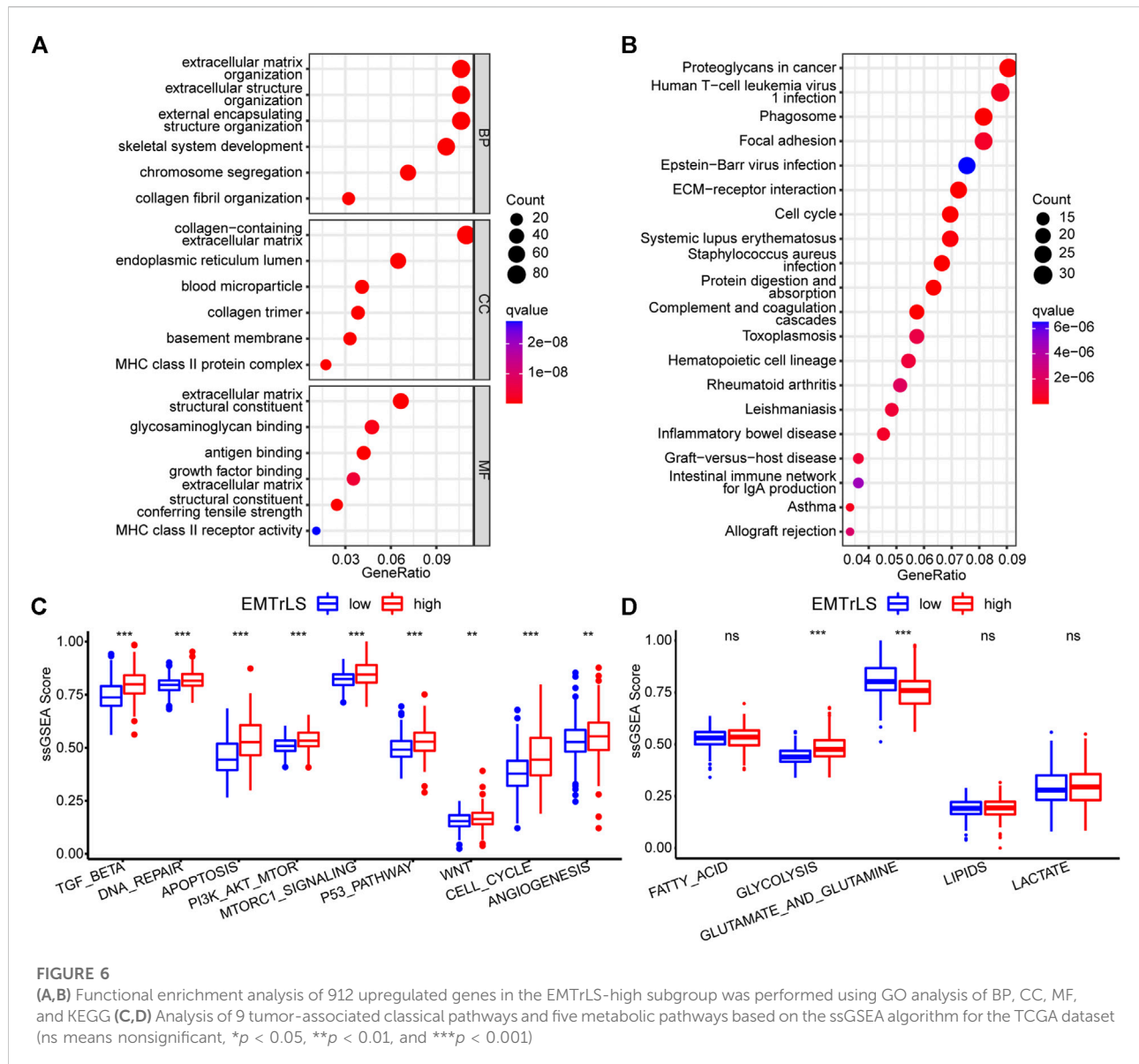
To clarify the functional differences associated with EMTrLS, we screened 1,252 differentially expressed genes between the EMTrLS-high and -low groups (Supplementary Table S6). GO and KEGG analyses were performed on 913 genes upregulated in the EMTrLS-high group. GO analysis revealed that pathways such as extracellular matrix, extracellular matrix structural constituent, and extracellular matrix organization were significantly upregulated. In addition, KEGG analysis highlighted that proteoglycans in cancer were upregulated in the EMTrLS-high group (Figures 6A,B). Interestingly, these genes upregulated in the EMTrLS-high group were also enriched in some immune-related pathways such as MHC class II protein completion and MHC class II receptor activity. This result suggested that there may be a potential association of EMTrLS with immune infiltration and immune function.

To further explore the tumor cell aggressiveness between the EMTrLS-high and -low groups, nine classical tumor pathways including the TGF $\beta$  pathway, DNA repair, apoptosis, PI3K-AKT-mTOR signaling, mTORC1 signaling, P53 pathway, WNT pathway, cell cycle, and angiogenesis were quantified by the ssGSEA algorithm. The scores of these pathways differed slightly across datasets but showed significant elevation in the EMTrLS-high group in at least two datasets. Collectively, EMTrLS characterized the malignant pathways associated with tumor development and metastasis (Figure 6C, Supplementary Figures S4A,B). In addition, we calculated the activation state of the metabolic pathways of LGG. The unlimited proliferative capacity and abnormal functional state of tumor cells often require a higher degree of glucose metabolism and protein metabolism to provide energy. The metabolic pathways of the glycolytic, glutamate and glutamine showed high activation in the EMTrLS-high group of all three datasets, indicating that EMTrLS is also related to the glycolytic pathways and the metabolic status of glutamate and glutamine in LGG (Figure 6D, Supplementary Figures S4C,D).

## Immune-related features

Given the interplay between EMT and the tumor immune microenvironment, we further analyzed the association between EMTrLS and immune infiltration in LGG. The stromal, immune, and ESTIMATE score were higher in the EMTrLS-high group, suggesting a higher level of immune infiltration (Figure 7A). By calculating ssGSEA scores for 28 immune cells, we found increased levels of immune cell infiltration, including CD8 T cells, central memory CD4 T cells, gamma delta T cells, B cells, Treg cells, T follicular helper cells, type 1 T helper cells, dendritic cells, macrophages, mast cells, MDSCs, NK cells, and NK T cells (Figure 7B, Supplementary Figures S4E–G). Correspondingly, we similarly found elevated immune infiltration scores in the EMTrLS-high group using the ImmuCellAI calculation. In addition, this method yielded 24 immune cells, of which T-cell cytotoxicity, exhausted T-cells, NK cells, Th17 cells, Tregs, Th1 cells, dendritic cells, monocytes, and macrophages were also significantly increased in the EMTrLS-high group (Figure 7C, Supplementary Figures S5A,B). Macrophages can activate and induce EMT through the TGF- $\beta$ , NF- $\kappa$ B pathway (Wei et al., 2019). This prompted us to further explore the correlation between EMT-related lncRNA and macrophage infiltration levels. The results showed that the Spearman correlation coefficients of EMTrLS scores with macrophages in the three datasets were 0.28, 0.33, and 0.26 ( $p < 0.001$ ) (Supplementary Figure S5C).

It was interesting to note that both analysis methods showed a significant increase in the infiltration of immune cells with tumor-killing effects, in addition to immune cells



such as macrophages, and Treg cells, which have the ability to suppress antitumor immunity, in the group with high EMTrLS scores. This result seems contradictory to the result that the EMTrLS-high group had a poorer prognosis. Therefore, we determined the antitumor immune differences between the different EMTrLS samples with the aid of the TIP algorithm. As a result, there was a significant increase in the third, sixth and seventh steps, which means that patients with low EMTrLS scores had a higher degree of T cell stimulation and activation, recognition, and killing of cancer cells (Figure 7D).

Furthermore, the differential expression of immune checkpoints and other molecules in the two LGG groups further explained the lower antitumor effect of the EMTrLS-

high group. BTLA, CD274, CD28, CD40, CD48, CD86, HAVCR2, ICOS, ICOSLG, IDO1, LAIR1, LGALS9, PDCD1, TNFRSF14, TNFRSF14, TNFRSF25, TNFRSF4, TNFSF4, and TNFRSF8 in the three databases were upregulated in the EMTrLS-high group. Activation of BTLA suppressed the function of CD8 cancer-specific T cells (Chen et al., 2019). The interaction of CD274 (PDL1) and PDCD1 (PD1) suppresses the cellular immune response of an organism, thus allowing tumor cells to evade surveillance and clearance by the immune system (Jiang et al., 2019). The increased expression of HAVCR2 (Bassez et al., 2021), LGALS9 (Wang et al., 2020), and LAIR1 (Peng DH et al., 2020) promoted the depletion and functional decline of T cells. In addition, ICOSLG

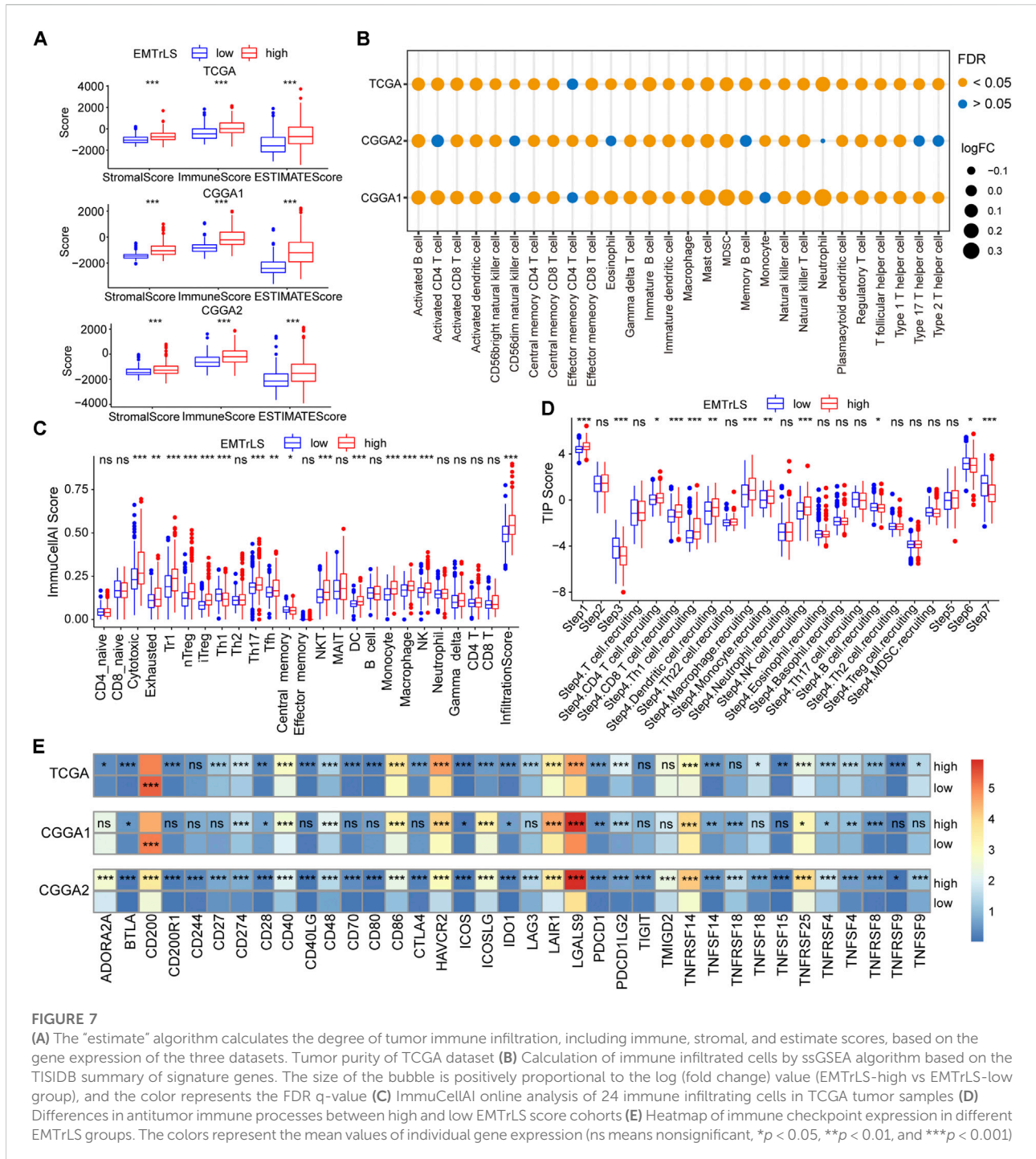


FIGURE 7

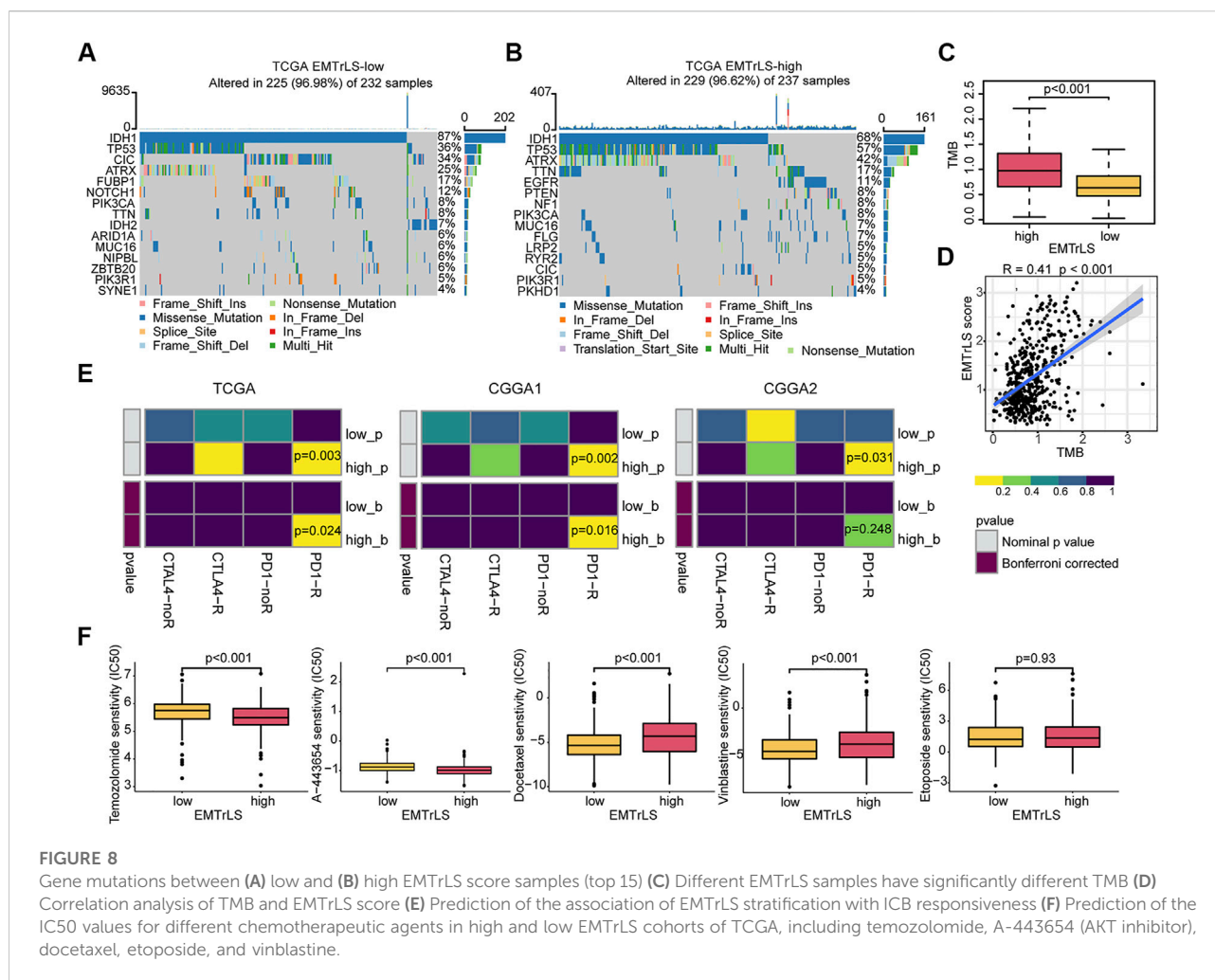
(A) The “estimate” algorithm calculates the degree of tumor immune infiltration, including immune, stromal, and estimate scores, based on the gene expression of the three datasets. Tumor purity of TCGA dataset (B) Calculation of immune infiltrated cells by ssGSEA algorithm based on the TISIDB summary of signature genes. The size of the bubble is positively proportional to the log (fold change) value (EMTrLS-high vs EMTrLS-low group), and the color represents the FDR q-value (C) ImmuneCellAI online analysis of 24 immune infiltrating cells in TCGA tumor samples (D) Differences in antitumor immune processes between high and low EMTrLS score cohorts (E) Heatmap of immune checkpoint expression in different EMTrLS groups. The colors represent the mean values of individual gene expression (ns means nonsignificant, \*p < 0.05, \*\*p < 0.01, and \*\*\*p < 0.001)

(Iwata et al., 2020), and IDO1 (Zhai et al., 2018) promoted Treg cell activity and suppressed antitumor immunity. CTLA-4, which was upregulated in the high EMTrLS group, exerted a negative regulation of the immune response upon binding to the equally upregulated T cell costimulatory ligand CD86 (Chen and Mellman, 2017) (Figure 7E, Supplementary Figures S6A–C).

## EMTrLS-related genomic features

Mutations in important genes have a role in patient prognosis and treatment outcome. We compared gene mutations between the two samples groups. As a result, the overall number of mutation frequencies was very similar; however, the specific mutated genes were very different. The





top 5 mutations with highest frequencies in samples with low EMTrLS scores were IDH1 (88%), TP53 (36%), CIC (34%), ARTX (25%), and FUBP1 (17%), while the top 5 mutations with highest frequencies in samples with high EMTrLS scores were IDH1 (68%), TP53 (57%), ATRX (42%), TTN (17%), and EGFR (11%) (Figures 8A,B). Increased mutations in TP53, an important regulator of cell growth, proliferation, and damage repair, often predict poor prognosis (Donehower et al., 2019). ATRX is a histone chaperone protein that loads histones onto telomeres and maintains heterochromatin (Qin et al., 2022). Loss of ATRX activates the immunosuppressive transcriptome and immune escape mechanisms in LGG cells.

Evidence from multiple sources suggests that higher TMB is associated with better outcomes after ICB treatment, particularly with PD-1 inhibitors (Chalmers et al., 2017). Stratified analysis showed that EMTrLS significantly differentiated the TMB values of LGG samples (Figure 8C). In addition, we found a positive correlation between TMB and EMTrLS score ( $\text{cor} = 0.41$ ,  $p < 0.001$ ) (Figure 8D). This result implied that there is a potential

correlation between EMTrLS scores and ICB immunotherapy outcomes that warrants further exploration.

## ICB responsiveness and chemotherapy outcomes based on EMTrLS score stratification

In clinical applications, ICB is a promising therapeutic approach, resulting in long-lasting survival benefits for patients with melanoma and hepatocellular carcinoma (Postow et al., 2015). With the submap algorithm, we predicted the effect of ICB treatment in patients with a different EMTrLS. As a result, LGGs with high EMTrLS were more likely to respond to anti-PD1 therapy (TCGA  $p = 0.003$ , Bonferroni-corrected  $p = 0.024$ ; CGGA1  $p = 0.003$ , Bonferroni-corrected  $p = 0.016$ ; CGGA2  $p = 0.031$ , Bonferroni-corrected  $p = 0.248$ ), suggesting that EMTrLS-based stratification has the potential to predict the response to anti-PD1 therapy (Figure 8E).

There is a significant difference in response to TMZ, which is a first-line chemotherapeutic agent administered after surgery in LGG patients (Omuro and DeAngelis, 2013). Accordingly, the “PRRophetic” package was employed to predict the response to chemotherapeutic agents among different EMTrLS groups. The results showed that the IC50 estimates were lower in the high-EMTrLS population ( $p < 0.001$ ) (Figure 8F, Supplementary Figures S6D,E). In addition, we included other therapeutic agents including A-443654 (Akt inhibitor), docetaxel, etoposide, and vinblastine. The results showed that the IC50 estimates of A-443654 were higher in the EMTrLS-low group in all three datasets ( $p < 0.001$ ). The IC50 estimates of docetaxel were lower in the EMTrLS-low group of the TCGA and CGGA1 datasets ( $p < 0.001$ ). This showed that the EMTrLS score can guide the clinical treatment of different LGG groups. And A-443654 and doxorubicin also have potential application in the treatment of glioma.

## Discussion

To our knowledge, this is the first application of machine learning to mine multiple datasets to identify the most well characterized EMT-related lncRNAs. In this study, we finally obtained three hub lncRNAs and used them to construct a scoring system to explore the characteristic patterns of EMT-related lncRNAs in LGG tumor cells. The stratified analysis revealed that samples with different EMTrLS possessed different clinicopathological parameters and EMT activation statuses. Another feature of the EMTrLS-high group was the activation of tumor malignancy-related pathways, as well as the activation of glycolytic pathways, and glutamate and glutamine metabolic pathways. In addition, EMTrLS distinguished the heterogeneity of immune infiltration and antitumor immunity in LGG samples. Notably, the EMTrLS-high subgroup was not only more sensitive to TMZ treatment but also more responsive to PD1 blockade therapy. Moreover, coupled with the powerful prognostic predictive power of the EMTrLS scoring system, it has the potential to guide more effective and precise treatment.

Our study revealed that abnormal states of EMT are present in LGG tumor tissues and correlated with malignant clinicopathological features, which may modulate the aggressiveness and treatment resistance of glioma cells (Pastushenko and Blanpain, 2019). Although, the study of genes associated with EMT is currently popular among researchers in the field of oncology. The functions of single lncRNAs such as TCL6 (Kulkarni et al., 2021), AP000695.4 (Liang et al., 2018), and RP11-390F4.3 (Peng PH et al., 2020), in the EMT process, have been discussed and confirmed in detail. Using bioinformatics approaches, many interesting EMT-related genes have been mined and predictive models have been constructed, including colorectal, bladder and clear cell renal cell carcinomas (Cao et al., 2020; Zhong et al., 2020; Yang et al.,

2021). The construction of these risk models reveals that EMT-related genes can be used as molecular markers for optimal tumor treatment and have powerful prognostic capabilities. However, there still exists a large number of lncRNAs that have not yet been tapped identified, resulting in incomplete understanding of EMT-related lncRNAs. Therefore more comprehensive work is needed to confirm the expression and function of EMT-related lncRNAs in LGG. This would be helpful for a deeper understanding of the molecular features and aggressiveness of low-grade gliomas.

In our study, a series of algorithms such as univariate Cox regression, RF, and LASSO regression, were used to identify the hub lncRNAs CRNDE, LINC00665, and NEAT1. The expression of CRNDE regulates cancer cell proliferation, migration, invasion, and apoptosis through multiple signaling pathways, such as WNT/ $\beta$ -catenin (Yu et al., 2017), PI3K/AKT (Liu et al., 2017), and mTOR (Wang et al., 2015) signaling pathways. LINC00665 is aberrantly expressed in a variety of human cancers and acts as an oncogene or tumor suppressor gene (Zhu et al., 2022). Silencing LINC00665, which is associated with poor prognosis in gliomas, suppressed the expression of markers of EMT (Lu et al., 2021). The last hub lncRNA NEAT1 promotes glioblastoma development and progression through the EGFR/NEAT1/EZH2/WNT/ $\beta$ -catenin axis (Chen et al., 2018). In addition, NEAT1 also promotes gliomagenesis through the mTOR signaling pathway (Yu et al., 2021). From these studies, it is clear that CRNDE, LINC00665, and NEAT1 are highly oncogenic and inextricably linked to EMT. This also explained that the signature models constructed from these three hub genes were able to describe the activation status of multiple tumor malignancy-related pathways. Importantly, no study has emerged that incorporates these three genes simultaneously in statistical model construction. Our constructed EMTrLS score demonstrated a stable correlation with EMT in multiple datasets and described the malignancy and aggressiveness of LGG with a strong predictive prognostic value. Of course, it cannot be said that these three pivotal lncRNAs can fully characterize all EMT-related lncRNAs. We hope that our study can help characterize EMT-related lncRNAs and provide a new way of thinking about tumorigenesis and development.

The characterization based on the three hub lncRNAs constructs was able to not only quantify the extent of EMT and malignant tumor pathway activation but also reveal its potential correlation with abnormal tumor metabolism. It is possible that dysregulation of WNT signaling control induces EMT, while increasing glucose consumption and lactate production through activation of pyruvate carboxylase (PC) gene expression (Lee et al., 2015). In addition to this, the hypoxic microenvironment of gliomas may be another cause of the relationship between glucose metabolism and EMT abnormalities. Hypoxia-inducible factors (HIFs) increase the

expression numerous glycolytic enzymes and can also activate EMT through various pathways, including TGF- $\beta$ , Notch, PI3K/AKT, WNT/ $\beta$ -catenin, and NF- $\kappa$ B (Tirpe et al., 2019).

The process of EMT stimulates the production of cytokines and chemokines in the TME to promote the infiltration of immune cells. In turn, TME is an effective inducer of EMT in tumor cells. The mutual maintenance of these two phenomena alters the expression and activity of various cell types that accumulate in the mesenchyme, particularly in response to various immune cell subtypes that influence tumor progression (Suarez-Carmona et al., 2017; Dongre and Weinberg, 2019). However, few studies have elucidated the potential mechanisms of EMT-related lncRNAs and immune infiltration in LGG. We applied multiple immunoassays to explore immune cell infiltration in LGG and determined the association with EMTrLS stratification. Interestingly, we observed a significant macrophage and Treg infiltration in the EMTrLS-high group, accompanied by an increased infiltration of immune cells with tumor-killing effects, such as CD4 T cells and CD8 T cells. It may seem paradoxical that there was a higher degree of infiltration of immune cells with antitumor function in EMTrLS-high samples with poor prognosis. However, some studies have shown that CD8 T cells and CD4 T cells, also lead to loss of E-cadherin expression in epithelial cells, accompanied by increased expression of vimentin and ZEB1 (EMT markers) (Santisteban et al., 2009; Goebel et al., 2015). Although the mechanism is unclear, cytokines such as IL-6 may facilitate this process (Chen et al., 2017). Furthermore, given the association of EMTrLS with EMT activation, it is not surprising that EMTrLS scores can be quantified to some extent for antitumor immune cells. Moreover, these immune cells that attack cancer cells are often overwhelmed by various immunosuppressive cells that also have increased infiltration in the EMTrLS-high group, such as MDSCs, macrophages, and Tregs. They can directly inhibit the antitumor function of T cells and NK cells and aid tumor progression (DeNardo and Ruffell, 2019; Togashi et al., 2019; Veglia et al., 2021). Notably, another stratification study of EMTrLS revealed that the EMTrLS-high group with poorer prognosis had overexpression of ICP, such as PD-1, PD-L1, CD86, HAVCR2, and ICOS. This further explains the reason why the antitumor immune function of the EMTrLS-high group was not well performed. The significant infiltration of immunosuppressive cells and overexpression of ICP may be the key mechanism of immune escape in EMTrLS-high samples.

TMB has been shown to be a reliable biomarker for tumor selection for ICB treatment. The exact mechanism is not known, but most patients who benefit from treatment tend to have a high TMB (Rizvi et al., 2015; Chan et al., 2019). In this study, we found a potential correlation between EMTrLS scores and TMB. EMTrLS stratified analysis indicated that the EMTrLS-high group had a higher TMB and may be a potential gaining group for immunotherapy. This is similar to

the results of our other study. A subclass mapping algorithm we used revealed that the EMTrLS-high group was more sensitive to PD1 blockade treatment. These results illustrate the potential of EMTrLS to guide more effective anti-PD1 therapy. However, due to the lack of LGG cohorts treated with ICB, the predictive power of EMTrLS requires further validation in more prospective trials. In predicting chemotherapy response, we found that the EMTrLS-high cohorts were more sensitive to TMZ and AKT inhibitor drugs. However, the mechanisms underlying the correlation between EMTrLS and chemotherapy response need to be further investigated. Likewise, a large number of drug response experiments are needed to confirm these predictions.

In conclusion, we constructed a novel EMTrLS with the help of machine algorithms, which can comprehensively assess the malignancy and prognosis of individual patients and provide new insights for diagnosis and clinical treatment decisions. However, it is undeniable that our study has some limitations. We should further confirm the specific mechanisms by which the three hub lncRNAs affect EMT through *in vivo* and *in vitro* experiments, and further validate their regulatory effects on tumor immune function and immune cells. In addition, our developed EMTrLS scoring system for the assessment of antitumor treatment sensitivity needs to be further verified in prospective studies and chemotherapy drug sensitivity assays. The EMTrLS model should include more clinical factors in practical clinical applications to improve predictive accuracy. We will incorporate these efforts into future studies.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

## Author contributions

SH, JZ, and XG conceived and designed the study. NW, JYD, and JZ collected the data. NW, JW, JD, HZ, BL, and SM provided analysis and technical support. JZ, XG, ZL, FW, and XY participated in the preparation of the figures. NW and JZ drafted the initial manuscript. SH, JYD and HJ critiqued and revised the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

This work was funded by the Zhejiang Provincial People's Hospital Talent Introduction Project (No. C-2021-QDJJ03-01).

## Acknowledgments

We sincerely appreciate the great work of TCGA and CGGA, as well as other sites such as UCSC, TISIDB, ImmuCellAI, and TIP for collating the data such that they are easily and quickly available to us.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462, 108–112. doi:10.1038/nature08460
- Bassez, A., Vos, H., Van Dyck, L., Floris, G., Arijis, I., Desmedt, C., et al. (2021). A single-cell map of intratumoral changes during anti-PD1 treatment of patients with breast cancer. *Nat. Med.* 27, 820–832. doi:10.1038/s41591-021-01323-8
- Cancer Genome Atlas Research Network (2015). Comprehensive, integrative genomic analysis of diffuse lower-grade gliomas. *N. Engl. J. Med. Overseas. Ed.* 372, 2481–2498. doi:10.1056/NEJMoal402121
- Cao, R., Yuan, L., Ma, B., Wang, G., Qiu, W., and Tian, Y. (2020). An EMT-related gene signature for the prognosis of human bladder cancer. *J. Cell. Mol. Med.* 24, 605–617. doi:10.1111/jcmm.14767
- Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., et al. (2017). Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* 9, 34. doi:10.1186/s13073-017-0424-2
- Chan, T. A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S. A., Stenzinger, A., et al. (2019). Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann. Oncol.* 30, 44–56. doi:10.1093/annonc/mdy495
- Chen, D. S., and Mellman, I. (2017). Elements of cancer immunity and the cancer-immune set point. *Nature* 541, 321–330. doi:10.1038/nature21349
- Chen, Q., Yang, D., Zong, H., Zhu, L., Wang, L., Wang, X., et al. (2017). Growth-induced stress enhances epithelial-mesenchymal transition induced by IL-6 in clear cell renal cell carcinoma via the Akt/GSK-3 $\beta$ / $\beta$ -catenin signaling pathway. *Oncogenesis* 6, e375. doi:10.1038/oncsis.2017.74
- Chen, Q., Cai, J., Wang, Q., Wang, Y., Liu, M., Yang, J., et al. (2018). Long noncoding RNA NEAT1, regulated by the EGFR pathway, contributes to glioblastoma progression through the WNT/ $\beta$ -Catenin pathway by scaffolding EZH2. *Clin. Cancer Res.* 24, 684–695. doi:10.1158/1078-0432.Ccr-17-0605
- Chen, Y. L., Lin, H. W., Chien, C. L., Lai, Y. L., Sun, W. Z., Chen, C. A., et al. (2019). BTLA blockade enhances Cancer therapy by inhibiting IL-6/IL-10-induced CD19(high) B lymphocytes. *J. Immunother. Cancer* 7, 313. doi:10.1186/s40425-019-0744-4
- DeNardo, D. G., and Ruffell, B. (2019). Macrophages as regulators of tumour immunity and immunotherapy. *Nat. Rev. Immunol.* 19, 369–382. doi:10.1038/s41577-019-0127-6
- Donehower, L. A., Soussi, T., Korkut, A., Liu, Y., Schultz, A., Cardenas, M., et al. (2019). Integrated analysis of TP53 gene and pathway alterations in the cancer Genome Atlas. *Cell Rep.* 28, 1370–1384. doi:10.1016/j.celrep.2019.07.001
- Dongre, A., and Weinberg, R. A. (2019). New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* 20, 69–84. doi:10.1038/s41580-018-0080-4
- Gittleman, H., Sloan, A. E., and Barnholtz-Sloan, J. S. (2020). An independently validated survival nomogram for lower-grade glioma. *Neuro. Oncol.* 22, 665–674. doi:10.1093/neuonc/noz191
- Goebel, L., Grage-Griebenow, E., Gorys, A., Helm, O., Genrich, G., Lenk, L., et al. (2015). CD4(+) T cells potently induce epithelial-mesenchymal-transition in premalignant and malignant pancreatic ductal epithelial cells—novel implications

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2022.942966/full#supplementary-material>

of CD4(+) T cells in pancreatic cancer development. *Oncoimmunology* 4, e1000083. doi:10.1080/2162402X.2014.1000083

Gritsch, S., Batchelor, T. T., and Gonzalez Castro, L. N. (2022). Diagnostic, therapeutic, and prognostic implications of the 2021 World Health Organization classification of tumors of the central nervous system. *Cancer* 128, 47–58. doi:10.1002/cncr.33918

Hoshida, Y., Brunet, J. P., Tamayo, P., Golub, T. R., and Mesirov, J. P. (2007). Subclass mapping: identifying common subtypes in independent disease data sets. *PLoS One* 2, e1195. doi:10.1371/journal.pone.0001195

Iwata, R., Hyoung Lee, J., Hayashi, M., Dianzani, U., Ofune, K., Maruyama, M., et al. (2020). ICOSLG-mediated regulatory T-cell expansion and IL-10 production promote progression of glioblastoma. *Neuro. Oncol.* 22, 333–344. doi:10.1093/neuonc/noz204

Jiang, X., Wang, J., Deng, X., Xiong, F., Ge, J., Xiang, B., et al. (2019). Role of the tumor microenvironment in PD-L1/PD-1-mediated tumor immune escape. *Mol. Cancer* 18, 10. doi:10.1186/s12943-018-0928-4

Kulkarni, P., Dasgupta, P., Hashimoto, Y., Shiina, M., Shahryari, V., Tabatabai, Z. L., et al. (2021). A lncRNA TCL6-miR-155 interaction regulates the src-akt-EMT network to mediate kidney cancer progression and metastasis. *Cancer Res.* 81, 1500–1512. doi:10.1158/0008-5472.CAN-20-0832

Lee, S. Y., Jeon, H. M., Ju, M. K., Jeong, E. K., Kim, C. H., Yoo, M. A., et al. (2015). Dlx-2 is implicated in TGF- $\beta$ - and Wnt-induced epithelial-mesenchymal, glycolytic switch, and mitochondrial repression by Snail activation. *Int. J. Oncol.* 46, 1768–1780. doi:10.3892/ijo.2015.2874

Li, T., Xie, J., Shen, C., Cheng, D., Shi, Y., Wu, Z., et al. (2015). Amplification of long noncoding RNA ZFAS1 promotes metastasis in hepatocellular carcinoma. *Cancer Res.* 75, 3181–3191. doi:10.1158/0008-5472.CAN-14-3721

Liang, H., Yu, T., Han, Y., Jiang, H., Wang, C., You, T., et al. (2018). LncRNA PTAR promotes EMT and invasion-metastasis in serous ovarian cancer by competitively binding miR-101-3p to regulate ZEB1 expression. *Mol. Cancer* 17, 119. doi:10.1186/s12943-018-0870-5

Liu, X. X., Xiong, H. P., Huang, J. S., Qi, K., and Xu, J. J. (2017). Highly expressed long non-coding RNA CRNDE promotes cell proliferation through PI3K/AKT signalling in non-small cell lung carcinoma. *Clin. Exp. Pharmacol. Physiol.* 44, 895–902. doi:10.1111/1440-1681.12780

Liu, J., Gao, L., Zhan, N., Xu, P., Yang, J., Yuan, F., et al. (2020). Hypoxia induced ferritin light chain (FTL) promoted epithelia mesenchymal transition and chemoresistance of glioma. *J. Exp. Clin. Cancer Res.* 39, 137. doi:10.1186/s13046-020-01641-8

Liu, S. J., Dang, H. X., Lim, D. A., Feng, F. Y., and Maher, C. A. (2021). Long noncoding RNAs in cancer metastasis. *Nat. Rev. Cancer* 21, 446–460. doi:10.1038/s41568-021-00353-1

Louis, D. N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W. K., et al. (2016). The 2016 World Health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathol.* 131, 803–820. doi:10.1007/s00401-016-1545-1

Lu, M., Qin, X., Zhou, Y., Li, G., Liu, Z., Geng, X., et al. (2021). Long non-coding RNA LINC00665 promotes gemcitabine resistance of Cholangiocarcinoma cells via



- regulating EMT and stemness properties through miR-424-5p/BCL9L axis. *Cell Death Dis.* 12, 72. doi:10.1038/s41419-020-03346-4
- Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28, 1747–1756. doi:10.1101/gr.239244.118
- Miao, Y. R., Zhang, Q., Lei, Q., Luo, M., Xie, G. Y., Wang, H., et al. (2020). ImmuCellAI: a unique method for comprehensive T-cell subsets abundance prediction and its application in cancer immunotherapy. *Adv. Sci.* 7, 1902880. doi:10.1002/advs.201902880
- Omuro, A., and DeAngelis, L. M. (2013). Glioblastoma and other malignant gliomas: a clinical review. *JAMA* 310, 1842–1850. doi:10.1001/jama.2013.280319
- Pastushenko, I., and Blanpain, C. (2019). EMT transition states during tumor progression and metastasis. *Trends Cell Biol.* 29, 212–226. doi:10.1016/j.tcb.2018.12.001
- Peng, D. H., Rodriguez, B. L., Diao, L., Chen, L., Wang, J., Byers, L. A., et al. (2020). Collagen promotes anti-PD-1/PD-L1 resistance in cancer through LAIR1-dependent CD8(+) T cell exhaustion. *Nat. Commun.* 11, 4520. doi:10.1038/s41467-020-18298-8
- Peng, P. H., Chieh-Yu Lai, J., Hsu, K. W., and Wu, K. J. (2020). Hypoxia-induced lncRNA RP11-390F4.3 promotes epithelial-mesenchymal transition (EMT) and metastasis through upregulating EMT regulators. *Cancer Lett.* 483, 35–45. doi:10.1016/j.canlet.2020.04.014
- Postow, M. A., Callahan, M. K., and Wolchok, J. D. (2015). Immune checkpoint blockade in cancer therapy. *J. Clin. Oncol.* 33, 1974–1982. doi:10.1200/JCO.2014.59.4358
- Qin, T., Mullan, B., Ravindran, R., Messinger, D., Siada, R., Cummings, J. R., et al. (2022). ATRX loss in glioma results in dysregulation of cell-cycle phase transition and ATM inhibitor radio-sensitization. *Cell Rep.* 38, 110216. doi:10.1016/j.celrep.2021.110216
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007
- Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., et al. (2015). Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348, 124–128. doi:10.1126/science.aaa1348
- Roh, W., Chen, P. L., Reuben, A., Spencer, C. N., Prieto, P. A., Miller, J. P., et al. (2017). Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci. Transl. Med.* 9, eaah3560. doi:10.1126/scitranslmed.aah3560
- Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., et al. (2019). TISIDB: an integrated repository portal for tumor-immune system interactions. *Bioinformatics* 35, 4200–4202. doi:10.1093/bioinformatics/btz210
- Santesteban, M., Reiman, J. M., Asiedu, M. K., Behrens, M. D., Nassar, A., Kalli, K. R., et al. (2009). Immune-induced epithelial to mesenchymal transition *in vivo* generates breast cancer stem cells. *Cancer Res.* 69, 2887–2895. doi:10.1158/0008-5472.CAN-08-3343
- Suarez-Carmona, M., Lesage, J., Cataldo, D., and Gilles, C. (2017). EMT and inflammation: inseparable actors of cancer progression. *Mol. Oncol.* 11, 805–823. doi:10.1002/1878-0261.12095
- Tirpe, A. A., Gulei, D., Ciortea, S. M., Crivii, C., and Berindan-Neagoe, I. (2019). Hypoxia: overview on hypoxia-mediated mechanisms with a focus on the role of HIF genes. *Int. J. Mol. Sci.* 20, E6140. doi:10.3390/ijms20246140
- Togashi, Y., Shitara, K., and Nishikawa, H. (2019). Regulatory T cells in cancer immunosuppression - implications for anticancer therapy. *Nat. Rev. Clin. Oncol.* 16, 356–371. doi:10.1038/s41571-019-0175-7
- Veglia, F., Sanseviero, E., and Gabrilovich, D. I. (2021). Myeloid-derived suppressor cells in the era of increasing myeloid cell diversity. *Nat. Rev. Immunol.* 21, 485–498. doi:10.1038/s41577-020-00490-y
- Wang, Y., Wang, Y., Li, J., Zhang, Y., Yin, H., and Han, B. (2015). CRNDE, a long-noncoding RNA, promotes glioma cell growth and invasion through mTOR signaling. *Cancer Lett.* 367, 122–128. doi:10.1016/j.canlet.2015.03.027
- Wang, H., Huo, X., Yang, X. R., He, J., Cheng, L., Wang, N., et al. (2017). STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4. *Mol. Cancer* 16, 136. doi:10.1186/s12943-017-0680-1
- Wang, L., Saci, A., Szabo, P. M., Chasalow, S. D., Castillo-Martin, M., Domingo-Domenech, J., et al. (2018). EMT- and stroma-related gene expression and resistance to PD-1 blockade in urothelial cancer. *Nat. Commun.* 9, 3503. doi:10.1038/s41467-018-05992-x
- Wang, M., Cai, Y., Peng, Y., Xu, B., Hui, W., and Jiang, Y. (2020). Exosomal LGALS9 in the cerebrospinal fluid of glioblastoma patients suppressed dendritic cell antigen presentation and cytotoxic T-cell immunity. *Cell Death Dis.* 11, 896. doi:10.1038/s41419-020-03042-3
- Wang, W., Hao, Y., Zhang, A., Yang, W., Wei, W., Wang, G., et al. (2021). miR-19a/b promote EMT and proliferation in glioma cells via SEPT7-AKT-NF- $\kappa$ B pathway. *Mol. Ther. Oncol.* 20, 290–305. doi:10.1016/j.omto.2021.01.005
- Wei, C., Yang, C., Wang, S., Shi, D., Zhang, C., Lin, X., et al. (2019). Crosstalk between cancer cells and tumor associated macrophages is required for mesenchymal circulating tumor cell-mediated colorectal cancer metastasis. *Mol. Cancer* 18, 64. doi:10.1186/s12943-019-0976-4
- Weller, M., van den Bent, M., Preusser, M., Le Rhun, E., Tonn, J. C., Minniti, G., et al. (2021). EANO guidelines on the diagnosis and treatment of diffuse gliomas of adulthood. *Nat. Rev. Clin. Oncol.* 18, 170–186. doi:10.1038/s41571-020-00447-z
- Xu, L., Deng, C., Pang, B., Zhang, X., Liu, W., Liao, G., et al. (2018). TIP: a web server for resolving tumor immunophenotype profiling. *Cancer Res.* 78, 6575–6580. doi:10.1158/0008-5472.CAN-18-0689
- Xu, A., Wang, X., Luo, J., Zhou, M., Yi, R., Huang, T., et al. (2021). Overexpressed P75CUX1 promotes EMT in glioma infiltration by activating beta-catenin. *Cell Death Dis.* 12, 157. doi:10.1038/s41419-021-03424-1
- Yang, J., Antin, P., Berx, G., Blanpain, C., Brabletz, T., Bronner, M., et al. (2020). Guidelines and definitions for research on epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* 21, 341–352. doi:10.1038/s41580-020-0237-9
- Yang, Y., Feng, M., Bai, L., Liao, W., Zhou, K., Zhang, M., et al. (2021). Comprehensive analysis of EMT-related genes and lncRNAs in the prognosis, immunity, and drug treatment of colorectal cancer. *J. Transl. Med.* 19, 391. doi:10.1186/s12967-021-03065-0
- Yoshihara, K., Shahmoradgoli, M., Martinez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Yu, B., Ye, X., Du, Q., Zhu, B., Zhai, Q., and Li, X. X. (2017). The long non-coding RNA CRNDE promotes colorectal carcinoma progression by competitively binding miR-217 with TCF7L2 and enhancing the wnt/ $\beta$ -catenin signaling pathway. *Cell. Physiol. Biochem.* 41, 2489–2502. doi:10.1159/000475941
- Yu, H., Xu, A., Wu, B., Wang, M., and Chen, Z. (2021). Long noncoding RNA NEAT1 promotes progression of glioma as a ceRNA by sponging miR-185-5p to stimulate DNMT1/mTOR signaling. *J. Cell. Physiol.* 236, 121–130. doi:10.1002/jcp.29644
- Zhai, L., Ladomersky, E., Lenzen, A., Nguyen, B., Patel, R., Lauing, K. L., et al. (2018). IDO1 in cancer: a gemini of immune checkpoints. *Cell. Mol. Immunol.* 15, 447–457. doi:10.1038/cmi.2017.143
- Zhong, W., Zhang, F., Huang, C., Lin, Y., and Huang, J. (2020). Identification of epithelial-mesenchymal transition-related lncRNA with prognosis and molecular subtypes in clear cell renal cell carcinoma. *Front. Oncol.* 10, 591254. doi:10.3389/fonc.2020.591254
- Zhu, J., Zhang, Y., Chen, X., Bian, Y., Li, J., and Wang, K. (2022). The emerging roles of LINC00665 in human cancers. *Front. Cell Dev. Biol.* 10, 839177. doi:10.3389/fcell.2022.839177