Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2023**; 28(10): 271
https://doi.org/10.31083/j.fbl2810271

Original Research

The EMT-Related Genes GALNT3 and OAS1 are Associated with Immune Cell Infiltration and Poor Prognosis in Lung Adenocarcinoma

Dan Luo[1,2,3,†], Mengying Fang[2,3,†], Le Shao[4,†], Jue Wang[1,†], Yuling Liang[2,3], Mengqin Chen[2,3], Xuemei Gui[2,3], Jie Yan[2,3], Wenjun Wang[2,3], Lili Yu[1,*], Xianming Fan[2,3,*], Qibiao Wu[1,5,*]

[1]Faculty of Chinese Medicine, State Key Laboratory of Quality Research in Chinese Medicine, Macau University of Science and Technology, Taipa, 999078 Macao, China
[2]Department of Respiratory and Critical Care Medicine, The Affiliated Hospital of Southwest Medical University, 646099 Luzhou, Sichuan, China
[3]Inflammation & Allergic Diseases Research Unit, The Affiliated Hospital of Southwest Medical University, 646099 Luzhou, Sichuan, China
[4]Center for Medical Research and Innovation, the First Hospital of Hunan University of Chinese Medicine, 410021 Changsha, Hunan, China
[5]Guangdong-Hong Kong-Macao Joint Laboratory for Contaminants Exposure and Health, Guangdong University of Technology, 510520 Guangzhou, Guangdong, China

*Correspondence: llyu@must.edu.mo (Lili Yu); fxm129@163.com (Xianming Fan); qbwu@must.edu.mo (Qibiao Wu)
†These authors contributed equally.

Academic Editor: Francesca Bruzzese
Submitted: 8 December 2022 Revised: 10 February 2023 Accepted: 20 February 2023 Published: 27 October 2023

Abstract

Background: Lung cancer is the main cause of cancer-related death, with epithelial-mesenchymal transition (EMT) playing an important role in the development of this disease. The EMT-related genes Polypeptide N-Acetylgalactosaminyltransferase 3 (*GALNT3*) and 2′-5′-Oligoadenylate Synthetase 1 (*OAS1*) are involved in numerous tumor processes. Although these genes have been extensively studied in cancer, they have yet to be analyzed by multi-omics in lung adenocarcinoma (LUAD). **Methods**: EMT-related genes were identified by R and Venn diagram. Cox regression and Kaplan-Meier analysis were performed to evaluate patient survival, and the Gene Expression Profiling Interactive Analysis (GEPIA) database was used for correlation analysis. GeneCards and R packages were used to explore gene characterization and functional annotation. The Tumor Immune Estimation Resource (TIMER), Human Protein Atlas (HPA), University of Alabama at Birmingham Cancer (UALCAN), and The Cancer Genome Atlas (TCGA) databases were used to investigate gene expression, which was then confirmed by RT-PCR. Clinicopathological analysis was carried out using the UALCAN database. Functional mechanisms and multi-omics analysis were performed using DNA Methylation Interactive Visualization Database (DNMIVD), Targetscan, TIMER, Tumor–immune System Interactions Database (TISIDB) and cBioportal. Diagnostic values were calculated using ROC curve analysis. **Results**: A total of 320 EMT-related genes were identified in LUAD. Their characteristics were confirmed in the Database for Annotation, Visualization and Integrated Discovery (DAVID) database by the intersection of 855 and 3600 different genes from the Gene Expression Omnibus (GEO) and EMTome databases, respectively. Expression of the EMT-related genes *GALNT3* and *OAS1* was associated with the prognosis of LUAD patients. A positive correlation was observed between the expression of *GALNT3* and *OAS1*, and their expression was higher in LUAD tissue than in normal lung tissue. This was confirmed using RT-PCR. Multi-omics analysis revealed that *GALNT3* and *OAS1* expression was associated with gene mutation and methylation, cellular immune infiltration, and several immune subtypes. A miRNA-*GALNT3/OAS1* regulatory network was also found. Receiver operating characteristic (ROC) curve analysis found that *GALNT3* and *OAS1* expression combined had superior diagnostic value to that of each marker alone. **Conclusions**: *GALNT3* and *OAS1* expression are associated with immune cell infiltration and poor prognosis in LUAD. Their combined expression has high diagnostic value; hence, *GALNT3* and *OAS1* may be valuable biomarkers for the early detection of LUAD.

Keywords: lung adenocarcinoma; EMT; immune infiltration; prognosis; *GALNT3*; *OAS1*

1. Introduction

Lung cancer is the leading cause of cancer-related death and affects the lives and health of millions of people worldwide [1]. It consists mainly of non-small-cell lung cancer (NSCLC) and small-cell lung cancer (SCLC), of which NSCLC accounts for approximately 85% of cases [2,3]. Lung adenocarcinoma (LUAD) is a histological subtype of NSCLC that accounts for about 50% of all new lung cancer cases [4,5]. Despite continued progress with various treatment options including surgery, targeted therapy, immunotherapy, radiotherapy, and chemotherapy, the prognosis of LUAD remains poor and the five-year survival rate is only about 20% [6–8]. Therefore, it is particularly important to identify effective biomarkers for the early diagnosis and accurate prognosis of LUAD.

Epithelial-mesenchymal transition (EMT) refers to the loss of epithelial characteristics and the acquisition of mesenchymal traits. The process of EMT is driven by a variety of factors, including loss of E-cadherin, transforming growth factor-β-driven receptor tyrosine kinase activation, and molecular transduction by downstream effectors [9–11]. EMT is an important biological process during embryonic development, cell differentiation and reprogramming, and cancer progression [12]. It can regulate multiple tumor functions such as integrated epigenetic modification, transcription control, alternative splicing, and protein stability [13,14]. Research on LUAD has demonstrated that EMT is an underlying cause of tumor development, metastasis and migration, tumor cell stemness, and tumor drug resistance [15–20]. However, few studies have explored the role of EMT-related genes in LUAD through multiple omics analyses.

The establishment of various biological databases and information collection has allowed researchers to conduct multi-omics research on tumors. To further explore the role of EMT-related genes in LUAD, we identified 320 EMT-related differentially expressed genes (EMT-related DEGs) from the GEO and EMTome databases. Two EMT-related genes with significant survival signatures, *GALNT3* and *OAS1*, were identified by Cox regression and Kaplan-Meier survival analysis. We subsequently analyzed immune cell infiltration, DNA methylation, gene mutation, miRNA regulation, prognostic significance, clinical relevance, and diagnostic value in relation to *GALNT3* and *OAS1* expression in LUAD. The mRNA expression of these potential biomarkers was also confirmed by RT-PCR.

2. Materials and Methods

2.1 Available Data from the GEO and EMTome Databases

RNA-seq data from 442 LUAD patients (including 442 primary tumor samples and 19 adjacent normal tissue samples) together with the associated clinical information was downloaded from the GSE68465 cohort (http://www.ncbi.nlm.nih.gov/geo/). EMT-related genes were downloaded from the EMTome database [21].

2.2 Screening and Functional Enrichment Analysis of EMT-Related Genes in LUAD

The expression of genes between tumor and normal lung tissue in the GSE68465 cohort was compared using the "limma" package in R. Those with a |log fold change| >2 and adjusted p-value < 0.05 were considered to be differentially expressed genes (DEGs). DEGs in the GEO database and EMT-related genes in the EMTome database were intersected to obtain EMT-related DEGs. The DAVID database was used for functional enrichment analysis of these EMT-related DEGs [22].

2.3 Prognostic Significance of EMT-Related DEGs in LUAD

Genes that were most strongly associated with prognosis were identified using the "survival" and "glement" packages in R. Univariate, multivariate, lasso Cox regression analysis and Kaplan-Meier survival analysis were performed to screen for prognostically significant EMT-associated genes. Correlation analysis of EMT-related DEGs was performed in GEPIA [23].

2.4 Genomic Characteristics, Expression Analysis, and Functional Enrichment Analysis of EMT-Related DEGs in LUAD

The genomic characteristics, including genomic and subcellular locations, of EMT-related DEGs with prognostic significance were first examined in GeneCards (https://www.genecards.org/) [24] and COMPARTMENTS (https://compartments.jensenlab.org/Search) [25]. Next, the differential expression of EMT-related DEGs with prognostic value was compared between tumor and normal lung tissue from the GSE68465 cohort using the "limma" package in R. Transcription levels for EMT-related DEGs with prognostic significance were then compared between pan-cancer and normal tissue in the TIMER database (https://cistrome.shinyapps.io/timer/) [26]. Protein expression levels of EMT-related DEGs with prognostic value in LUAD were also investigated in the Human Protein Atlas (HPA) database (http://www.proteinatlas.org). EMT-related genes were grouped according to high or low expression in the GSE68465 cohort, with the "clusterProfiler" package in R then used to perform functional enrichment analysis.

2.5 Correlation of EMT-Related DEG Expression with Clinical Characteristics in LUAD

Univariate and multivariate Cox regression analyses were used to evaluate correlations between EMT-related DEG expression and various clinicopathological features in LUAD patients, including age, pathological T stage, *TP53* mutation status, and pathological N stage. The prognostic significance of EMT-related DEG expression in clinicopathological subgroups of LUAD patients was also investigated in the UALCAN database [27].

2.6 DNA Methylation Status and miRNA Regulation of EMT-Related DEGs in LUAD

The DNA methylation level and prognostic significance of EMT-related DEGs in LUAD were evaluated in the DNMIVD database [28]. Online databases were used to predict the regulatory relationship between target genes and miRNAs. The TargetScan database (https://www.targetscan.org/vert_80/) [29], miRDB database (http://www.miRDB.org) [30], and miRTarBase database (https://mirtarbase.cuhk.edu.cn/) [31] were used to predict potential miRNAs upstream of EMT-related DEGs. The predicted miRNAs from the three databases were then intersected to identify miRNAs that could regulate the expression of EMT-related DEGs. Finally, the relationship between EMT-related DEGs and the miRNA expression level was further evaluated using the starBase database [32].

2.7 Correlation of EMT-Related DEG Expression with Immune Cell Infiltration in LUAD

The TIMER database was used to correlate the expression of EMT-related DEGs with tumor-infiltrating immune cells including CD4+ T cells, CD8+ T cells, macrophages, dendritic cells, B cells, and neutrophils. Immune cell markers were obtained from the website of R&D Systems (www.rndsystems.com/cn/resources/cell-markers/immune-cells) and analyzed in TIMER and GEPIA. The associations between EMT-associated DEGs and immune cell subtypes, tumor infiltrates (TILs), immune stimulants, immunosuppressants, and major histocompatibility class (MHC) molecules in LUAD were examined in the TISIDB database (cis.hku.hk/TISIDB) [33]. Kaplan–Meier analysis was used to investigate the prognostic significance of immune cell infiltration (https://kmplot.com/analysis/) [34].

2.8 DNA Mutation of EMT-Related DEGs in LUAD and Confirmation of Expression Levels

The cBioPortal database (http://cbioportal.org) was used to examine the DNA mutation status of EMT-related DEGs in 586 LUAD samples from the TCGA database [35]. We explored mutation sites in EMT-related DEGs and assessed gene mutation co-occurrence patterns between EMT-related DEG signatures and other proteins in LUAD. The top 10 altered genes that co-occurred with mutated genes were selected according to the percentage of alteration. In addition, the mRNA expression level of EMT-related DEGs was confirmed in TCGA-LUAD samples from the UALCAN database.

2.9 Evaluation of the Diagnostic Accuracy of EMT-Related DEGs in LUAD

The OmicStudio tool (https://www.omicstudio.cn/tool) was employed to construct ROC curves in order to assess the diagnostic value of EMT-related genes. A single gene or a combination of EMT-related genes was used to assess diagnostic value.

2.10 Cell Culture

Human LUAD (H1975) and bronchial epithelial (BEAS-2B) cell lines were purchased from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). These cells had been performed by STR analysis and their DNA profiles were consistent with the recorded. Additionally, mycoplasma testing was conducted on the cell lines, and the result is negative. BEAS-2B cells were grown in DMEM (Bio-Channel, Nanjing, Jiangsu, China), while H1975 cells were grown in ERPMI-1640 medium (Bio-Channel, Nanjing, Jiangsu, China). All cells were supplemented with penicillin (100 U/mL), streptomycin (100 U/mL) and 10% heat-inactivated fetal bovine serum (Transgen, Beijing, China) and cultured in a humidified environment at 37 °C with 5% CO_2.

2.11 RNA Isolation and RT-PCR

Adherent cells were used in the logarithmic growth phase, which corresponded to a cell density of approximately 80%. TRIzol reagent (Invitrogen, CA, USA) was used to extract total RNA from the cells and transcription reagent (Transgen, Beijing, China) was then used to transcribe the RNA into cDNA. RT-PCR was performed in the 7500 Real-Time PCR System (BIO-RAD) using SYBR qPCR mix (Transgen, Beijing, China) and GAPDH as the endogenous reference [36]. The primer sequences used were: GAPDH, 5′-GTCTCCTCTGACTTCAACAGCG-3′ (forward) and 5′-ACCACCCTGTTGCTGTAGCCAA-3′ (reverse); *GALNT3*, 5′-GCTGCAGTTTCATGTTAGGG-3′ (forward) and 5′-ACAGCCCCAATAACCGTATGAA-3′ (reverse); *OAS1*, 5′-GCTGGCTGAAAGCAACAGTG-3′ (forward) and 5′-TCCAGTCCTCTTCTGCCTGT-3′ (reverse). The relative initial copy number of the target gene was calculated using the $2^{-\triangle\triangle CT}$ method.

2.12 Statistical Analysis

R 4.1.1 (R studio Version 1.4.1717) and GraphPad Prism 8.0 (San Diego, California USA, www.graphpad.com) were used for statistical analysis. The Wilcoxon test was used to screen for genes that were differentially expressed. Prognostic factors were assessed using univariate, Lasso, and multivariate Cox regression analysis. The Kaplan–Meier method was used to evaluate the correlation between gene expression level and overall survival rate. The student's *t*-test was used to analyze the RT-PCR results. The 95% confidence interval (CI) and hazard ratio (HR) was used to calculate overall survival. $p < 0.05$ was used as the threshold for statistical significance.

3. Results

3.1 Identification and Functional Analysis of EMT-Related Genes in LUAD from the GSE68465 Cohort and Using the EMTome Database

The flow chart for this study is shown in Fig. 1A. Microarray-derived expression database for the GSE68465 LUAD cohort was screened by "limma" in R (adjusted *p*-value < 0.05 and $|\text{logFC}| > 2$). A total of 855 DEGs were obtained, comprising 650 upregulated genes and 205 downregulated genes (Fig. 1B). The EMTome database is a resource that includes pan-cancer analysis of EMT genes and their associated characteristics. EMT signature genes were downloaded from the EMTome database. A total of 320 EMT-related DEGs were obtained by the intersection of 855 DEGs from the GEO database with 3600 EMT signature genes from the EMTome database (Fig. 1C). Subsequently, functional enrichment analysis and enrichment pathway analysis was performed on these EMT-related genes in the DAVID database. As shown in **Supplementary Fig. 1A–C**, EMT-related DEGs were mainly enriched in extracellular matrix organization via Biological Process (BP), extracellular space via Cellular Component (CC), and heparin-

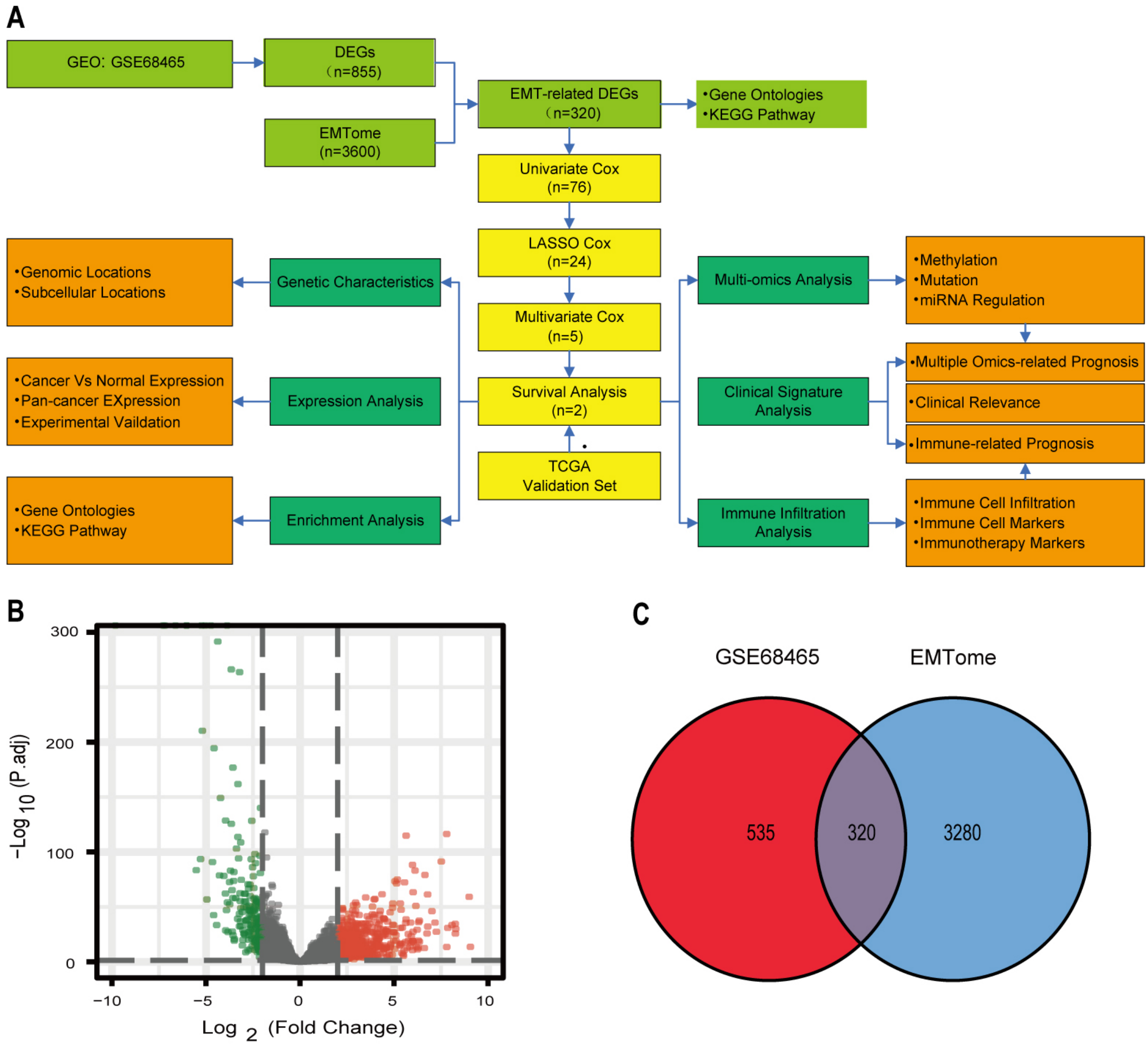

Fig. 1. Flowchart for identifying EMT-related genes in LUAD using the GES68465 cohort and the EMTome database. (A) Flowchart used for data collection and the methods used in this study. (B) Volcano plots of DEGs in LUAD from the GSE68465 cohort. (C) Venn diagram of EMT-related genes in LUAD from the GSE68465 cohort and EMTome database. GEO, Gene Expression Omnibus; DEGs, differentially expressed genes; EMT, epithelial-mesenchymal transition; KEGG, Kyoto Encyclopedia of Genes and Genomes; TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma.

binding via Molecular Function (MF). The significant signaling pathways for EMT-related DEGs were mainly enriched in amoebiasis, the PI3K-Akt signaling pathway, and ECM-receptor interactions (**Supplementary Fig. 1**).

3.2 Cox Regression and Kaplan-Meier Survival Analysis of EMT-Related Genes in LUAD

The "survival" package in R was used to perform univariate Cox regression analysis on 320 EMT-related DEGs. This identified 76 EMT-related genes with significant prognostic value. Lasso Cox analysis using the "glement" package in R identified 24 genes (Fig. 2A,B). Multivariate Cox analysis was used for further gene screening. A forest plot of hazard ratio (HR) (Fig. 2C) revealed that 5 genes had adverse prognostic significance (*GALNT3*, *MUC5AC*, *OAS1*, *TMPRSS11E*, *VSIG4*). The "survival" package in R was then used to screen for the strongest prognostic gene signatures. Kaplan–Meier analysis showed that GALNT3 and OAS1 were associated with significantly worse prognosis of LUAD patients in the GSE68465 cohort (Fig. 2D–H). Correlation analysis using the GEPIA database also revealed that *GALNT3* and *OAS1* expressions were closely correlated (r = 0.46, p-value = 1.3 $\times$ 10^{-44}) (Fig. 2I).

3.3 Genomic Characteristics, Differential Expression, and Functional Enrichment of GALNT3 and OAS1 in LUAD

The genomic location of *GALNT3* (via GeneCards) was q24.3 on chromosome 2 (Fig. 3A), while its subcellular location (via COMPARTMENTS) was mainly in the Golgi apparatus and extracellular space (Fig. 3C). The genomic location of *OAS1* was q24.13 on chromosome 12 (Fig. 3B), while its subcellular location was centrally in

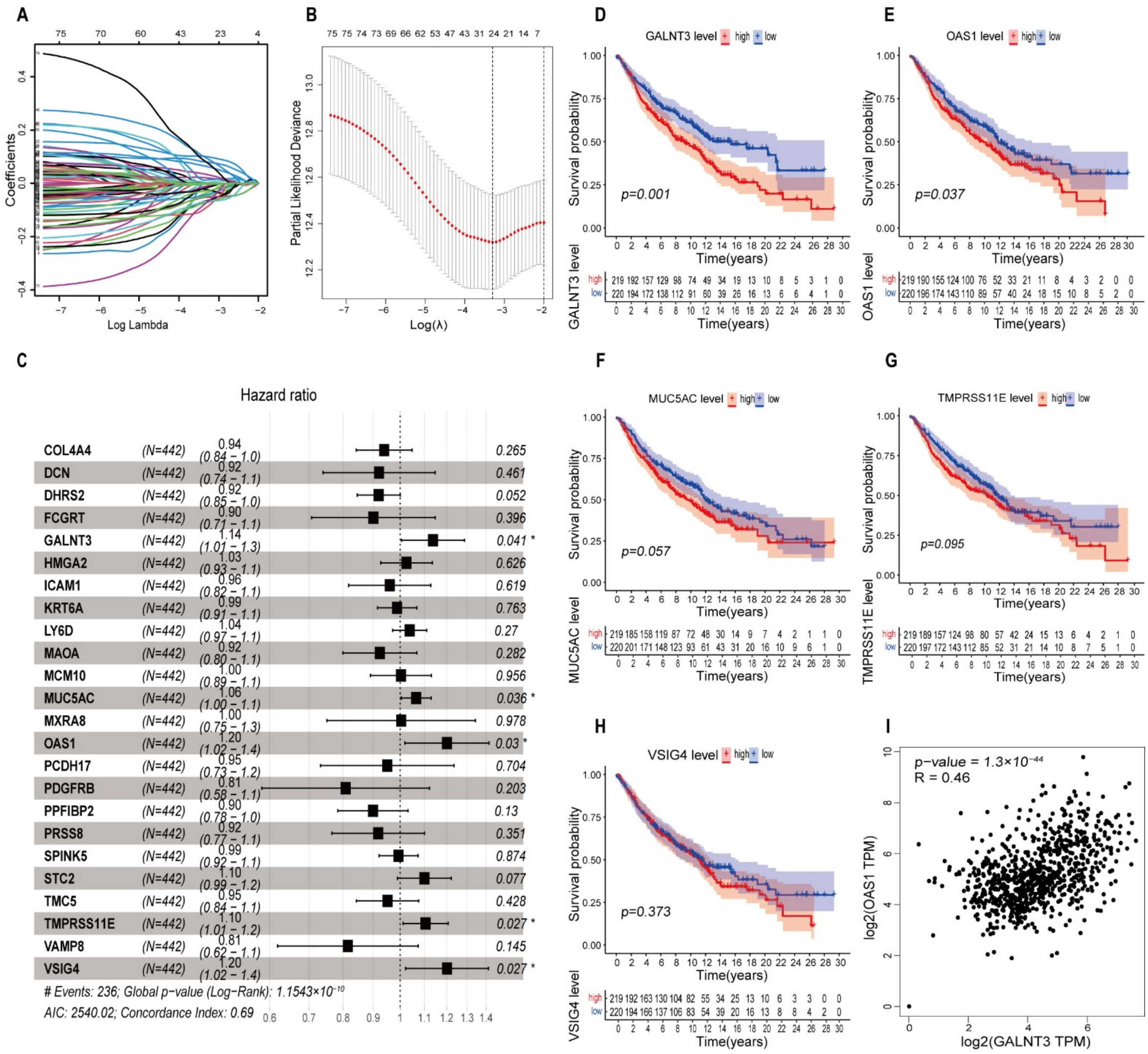

Fig. 2. Cox regression and Kaplan-Meier analysis were used to identify EMT-related genes with prognostic significance in LUAD. (A) The coefficient profiles of lasso regression analysis. (B) Minimal criteria were used to choose variables in lasso regression. (C) Forest plot showing the results of multivariate Cox regression analysis. (D–H) Kaplan–Meier survival analysis of GSE68465 patients according to the expression of *GALNT3*, *OAS1*, *MUC5AC*, *TMPRSS11E*, and *VSIG4*. (I) Correlation of *GALNT3* and *OAS1* expression in LUAD.

the nucleus and cytosol (Fig. 3D). *GALNT3* and *OAS1* mRNA were overexpressed in LUAD tumor tissue from the GSE68465 cohort compared to normal tissue (Fig. 3E,H). *GALNT3* and *OAS1* protein expression were also higher in LUAD tumor tissue from the HPA database compared to normal lung tissue (Fig. 3F,I). The differential expression of *GALNT3* and *OAS1* between tumor and normal tissue in all TCGA tumors was investigated using the TIMER database (Fig. 3G,J). This analysis indicated that *GALNT3* and *OAS1* were overexpressed in most tumors. As shown in **Supplementary Fig. 2**, Gene Ontology (GO) annotations revealed that *GALNT3* was mainly enriched in skin development (**Supplementary Fig. 2A**), while Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis showed that PI3K/Akt signaling was the most enriched pathway (**Supplementary Fig. 2B**). *OAS1* was mainly enriched in response to the virus (**Supplementary Fig. 2C**), with Hepatitis C being the most enriched pathway according to KEGG pathway analysis (**Supplementary Fig. 2D**).

3.4 Prognostic Significance of GALNT3 and OAS1 Expression and Clinicopathological Characteristics in LUAD

Univariate and multivariate Cox regression analyses were used to study the relationships between *GALNT3* and *OAS1* mRNA expression and the prognosis and clinical characteristics of LUAD. Univariate Cox analysis showed that *GALNT3* (HR = 1.230; 95% confidence interval (CI) = 1.097–1.378; $p < 0.001$) and *OAS1* (HR = 1.201; 95% CI = 1.048–1.377; $p = 0.008$) mRNA expression levels were significantly associated with poor prognosis of LUAD in the GSE68465 cohort. The clinicopathological factors of age, gender, T stage and N stage were also associated with poor

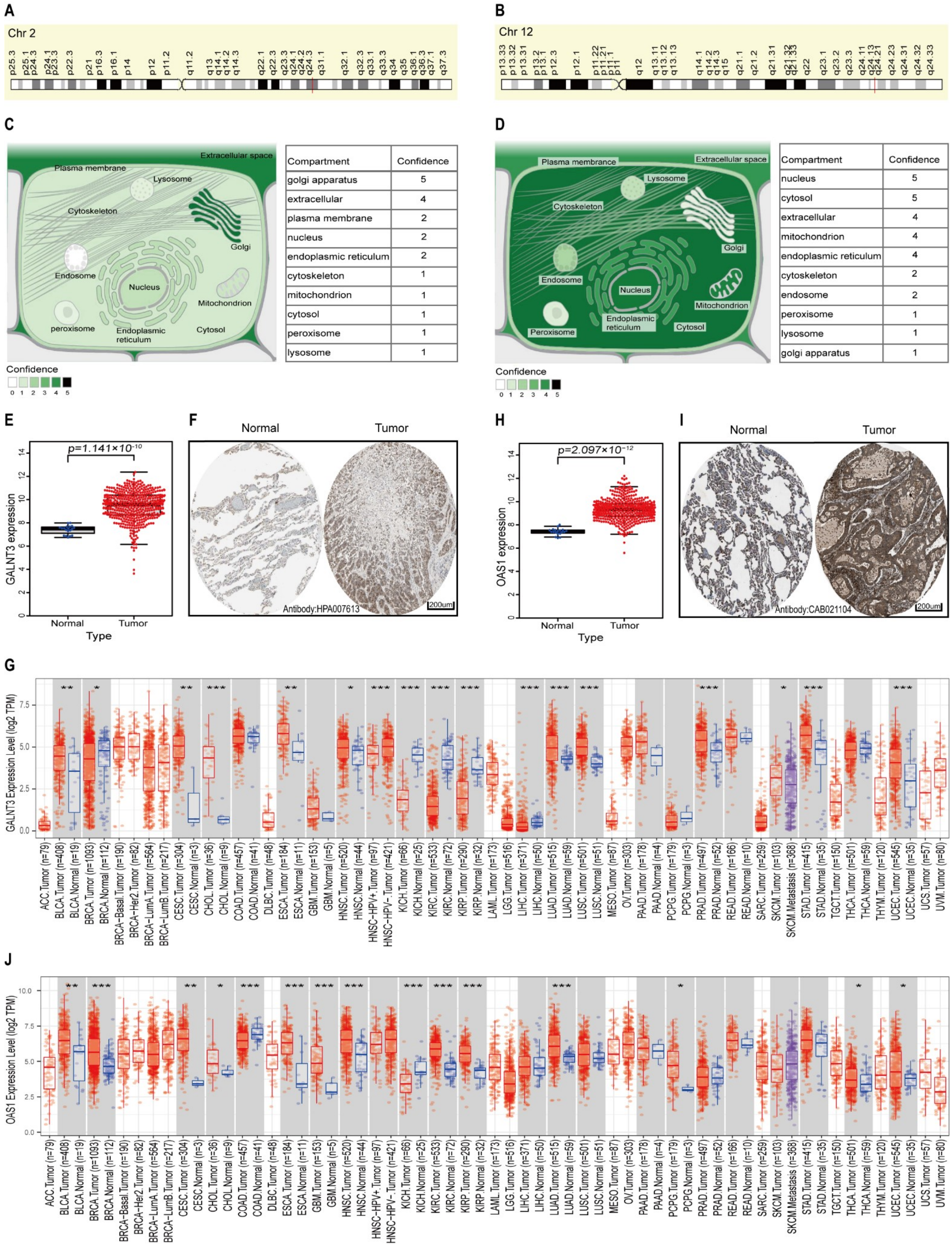

Fig. 3. Genomic characteristics and differential expression of *GALNT3* and *OAS1* in LUAD. (A,B) Genomic locations of *GALNT3* and *OAS1* as viewed by GeneCards. (C,D) Subcellular locations of *GALNT*3 and *OAS1* as viewed by COMPARTMENTS. (E) The mRNA expression level of *GALNT3* in the GSE68465 cohort. (F) Immunohistochemical staining for *GALNT3* in normal lung tissue and LUAD tissue from the HPA database. (G) Differential expression of *GALNT3* between all tumors and adjacent normal tissue in TCGA via the TIMER database. (H) The mRNA expression level of *OAS1* in the GSE68465 cohort. (I) Immunohistochemical staining for *OAS1* in normal lung tissue and in LUAD tissue from the HPA database. (J) Differential expression of *OAS1* between all tumors and adjacent normal tissue in TCGA via the TIMER database (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$). HPA, Human Protein Atlas; TIMER, Tumor Immune Estimation Resource.

survival (Fig. 4A,B). Multivariate Cox regression analysis showed that high mRNA expression of *GALNT3* was independently associated with poor prognosis (HR = 1.187; 95% CI = 1.055–1.334; p = 0.004), but not *OAS1* (HR = 1.094; 95% CI = 0.953–1.2561; p = 0.203) (Fig. 4C,D). These results imply that *GALNT3* expression is an independent biomarker of poor prognosis in LUAD. Although *OAS1* expression showed no independent prognostic significance in multivariate analysis, this may be due to other confounding factors. Therefore, we comprehensively evaluated *OAS1* as a marker of poor prognosis in LUAD by adding a multifaceted analysis method and database.

We further explored the prognostic significance of *GALNT3* and *OAS1* expression levels in the UALCAN database, as well as their correlation with clinical features. The mRNA expression levels for both genes correlated closely with stage, *TP53* mutation status and pathological N stage (Fig. 4E–J). Although the association with the pathological N3 stage was not statistically significant, this may have been due to the small number of samples. Patients with high expression of *GALNT3* and *OAS1* showed poor survival (Fig. 4K,L). Together, the above results indicate that high expression levels of *GALNT3* and *OAS1* may be associated with poor prognosis of LUAD patients.

3.5 Potential Molecular Mechanisms Involving GALNT3 and OAS1 in LUAD

DNA methylation status is strongly related to gene dysregulation in cancer cells. We explored the methylation status of promoter regions in *GALNT3* and *OAS1* in LUAD samples from the DNMIVD database. As shown in Fig. 5A,B, both the *GALNT3* (p-value = 1.95 $\times$ 10^{-3}) and *OAS1* (p-value = 3.09 $\times$ 10^{-25}) genes were significantly demethylated in LUAD tissue compared with normal lung tissue. Next, the samples were grouped into low, middle, and high methylation groups, with DNA methylation beta values of 0.3 and 0.7 used as cutoffs. The log-rank test was used then to calculate the p-value for prognostic significance of these groups. Kaplan–Meier survival analysis showed that the DNA methylation status of *GALNT3* was closely related to prognosis (p-value = 4.02 $\times$ 10^{-3}), but not that of *OAS1* (p-value = 0.941) (Fig. 5C,D). Since the expression of oncogenes is regulated by miRNAs, potential upstream miRNAs that regulate *GALNT3* and *OAS1* mRNA expression were predicted using the TargetScan, miRDB, and miRTarBase databases. These online databases identified 27 overlapping miRNAs that target *GALNT3*. The StarBase database was used to analyze the coexpression of *GALNT3* with the 27 overlapping miRNAs in LUAD. This revealed the mRNA expression level of *GALNT3* was negatively associated with the expression of hsa-miR-122-5p (r = –0.092, p value = 3.75 $\times$ 10^{-2}) (Fig. 5E). Only two overlapping miRNAs that target *OAS1* were identified, and neither was co-expressed with *OAS1* (Fig. 5F).

3.6 Correlation of GALNT3 and OAS1 Expression with Immune Markers in LUAD

Previous studies have shown that EMT-related genes play a critical role in tumor immunity. Here, we analyzed the correlation of *GALNT3* and *OAS1* expression with immune cell infiltration in LUAD. High *GALNT3* expression was associated with high levels of many immune cell populations, including B cells (partial.cor = –0.129, p = 4.40 $\times$ 10^{-3}), CD8+ T cells (partial.cor = 0.148, p = 1.02 $\times$ 10^{-3}), macrophages (partial.cor = 0.119, p = 8.47 $\times$ 10^{-3}), neutrophils (partial.cor = 0.162, p = 3.59 $\times$ 10^{-4}), and dendritic cells (partial.cor = 0.165, p = 2.59 $\times$ 10^{-4}). Tumor purity (partial.cor = –0.055, p = 2.20 $\times$ 10^{-1}) and CD4+ cells (partial.cor = 0.047, p = 3.05 $\times$ 10^{-1}) were not significantly associated with *GALNT3* expression (Fig. 6A). High *OAS1* expression was correlated with tumor purity (partial.cor = –0.154, p = 5.92 $\times$ 10^{-4}), neutrophils (partial.cor = 0.235, p = 1.85 $\times$ 10^{-7}), and dendritic cells (partial.cor = 0.159, p = 4.19 $\times$ 10^{-4}), but not with B cells (partial.cor = –0.08, p = 7.90 $\times$ 10^{-2}), CD4+ cells (partial.cor = –0.08, p = 7.90 $\times$ 10^{-2}), CD8+ T cells (partial.cor = 0.067, p = 1.38 $\times$ 10^{-1}) or macrophages (partial.cor = 0.062, p = 1.70 $\times$ 10^{-1}) (Fig. 6B).

We further analyzed the correlations between *GALNT3* and *OAS1* expression in LUAD with immune cell infiltration using the TISIDB database. Samples were divided into TGF-b dominant, immunologically quiet, inflammatory, IFN-gamma dominant, and wound healing groups. *GALNT3* expression was highest in the lymphocyte-depleted group and lowest in the inflammatory cell group (Fig. 6C), while *OAS1* expression was highest in the IFN-gamma-dominant cell group and weakest in the inflammatory cell group (Fig. 6D). The relationships between *GALNT3* and *OAS1* expression in LUAD with TILs, MHC molecules, immune inhibitors, and immune stimulators are shown in **Supplementary Fig. 3**.

To confirm the associations between *GALNT3* and *OAS1* expression in LUAD with immune infiltrating cells, further analysis was conducted using the TIMER and GEPIA databases (**Supplementary Tables 1,2**). The results confirmed that *GALNT3* and *OAS1* expression levels were closely associated with those of Th2 cell immune marker genes, including *STAT6*, *CCR8*, and *HAVCR1*. Numerous factors could modulate the prognostic significance of *GALNT3* and *OAS1* expression in LUAD. We hypothesized that immune cell infiltration could affect their prognostic impact in LUAD. Therefore, we further analyzed the prognostic significance of *GALNT3* and *OAS1* in the Kaplan–Meier plotter database according to the presence of immune cell infiltration. This showed that high *GALNT3* expression was most strongly associated with poor prognosis in LUAD cases that had high levels of macrophage, regulatory T cell and type 2 T helper cell infiltration (Fig. 6E–G). Similarly, high *OAS1* expression was associated with poor prognosis in cases that showed strong infiltration with

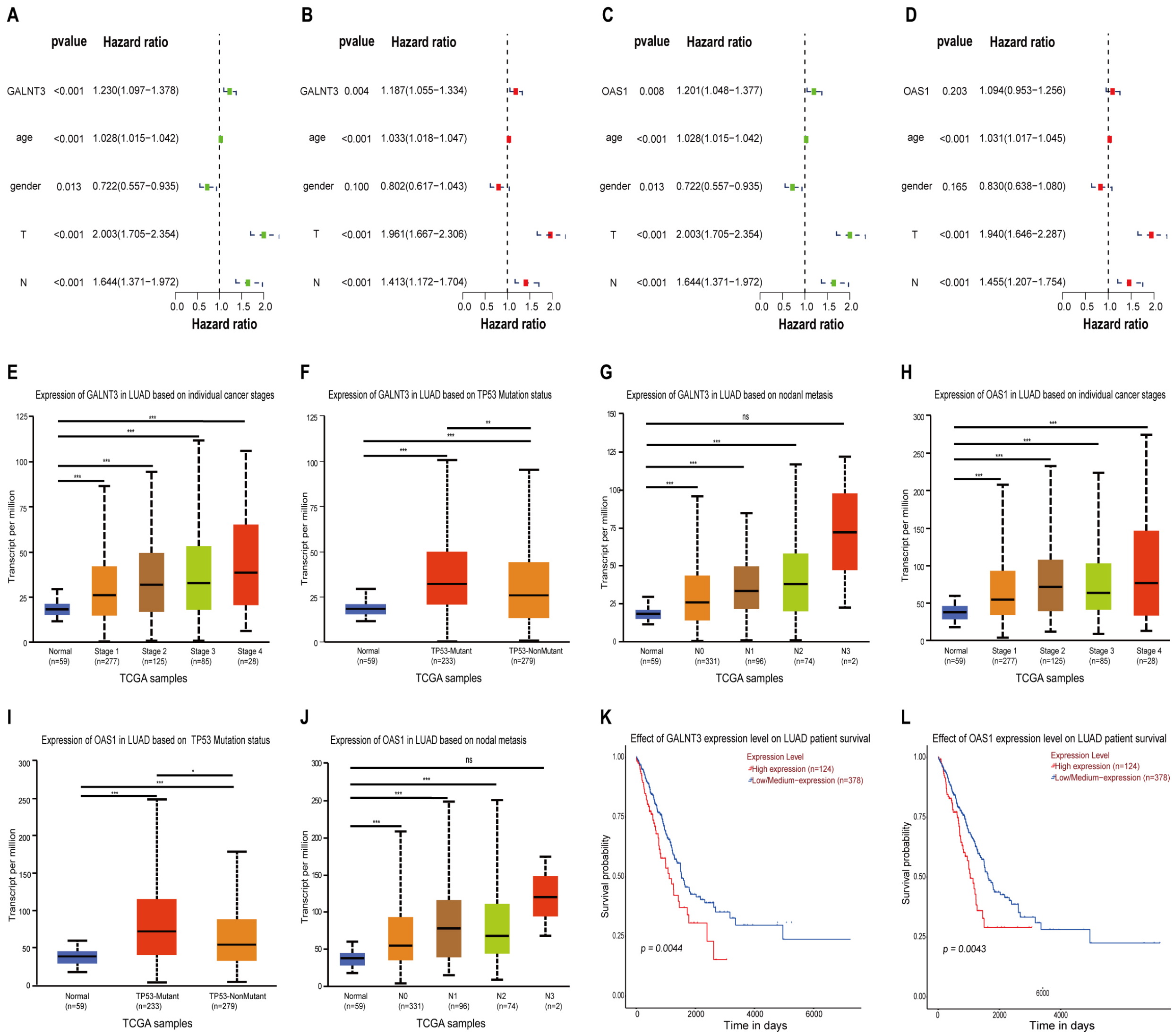

Fig. 4. Prognostic significance of *GALNT3* and *OAS1* expression and associations with clinicopathological features in LUAD. (A) Univariate and (B) multivariate Cox analyses showing independent prognostic significance for *GALNT3* expression in LUAD from the GSE68465 cohort. (C) Univariate and (D) multivariate Cox analyses showing the prognostic significance of *OAS1* expression in LUAD from the GSE68465 cohort. (E–G) Correlation of *GALNT3* expression in LUAD with individual cancer stages, *TP53* mutation status and nodal metastasis status. (H–J) Correlation of *OAS1* expression in LUAD with individual cancer stages, *TP53* mutation status, and nodal metastasis statuses. (K,L) Kaplan-Meier analysis showing the prognostic significance of *GALNT3* and *OAS1* expression for the survival of LUAD patients (ns, not statistically significant, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

CD8+ T cells, macrophages, mesenchymal stem cells, and type 2 T helper cells (Fig. 6H–K). These results suggest that LUAD with high *GALNT3* and *OAS1* expression may potentially respond to immunotherapy through the regulation of immune cell infiltration.

3.7 DNA Mutation, Confirmation of mRNA Expression, and the Diagnostic Value of GALNT3 and OAS1 in LUAD

We next examined the frequencies and types of genetic alterations in *GALNT3* and *OAS1* in TCGA-LUAD using the cBioPortal database. The genetic alterations found in *GALNT3* were amplification and mutations. Somatic mutations in *GALNT3* were found in 2.6% of cases, with the mutation sites located within amino acids 0 and 633 and in the Glycos-transf-2 and Ricin-B-lectin domains. Four missense mutations were reported, with one being the hotspot mutation *E545K* (Fig. 7A,C). Gene alterations in *OAS1* consisted of amplification, mutations and deletions. Somatic mutations were found in 3% of cases, with mutation sites in the 0 to 400 amino acid NTP-transf-2 and OAS1-C domains. Three missense mutations were reported, with one being the hotspot mutation *T251K* (Fig. 7B,C). Genetic alterations in *GALNT3* and *OAS1* correlated with the frequency and pattern of genetic alterations in several other genes (Fig. 7D,E). The top 10 co-occurring genes with *GALNT3* alteration were *TTN* (47% co-alteration fre-

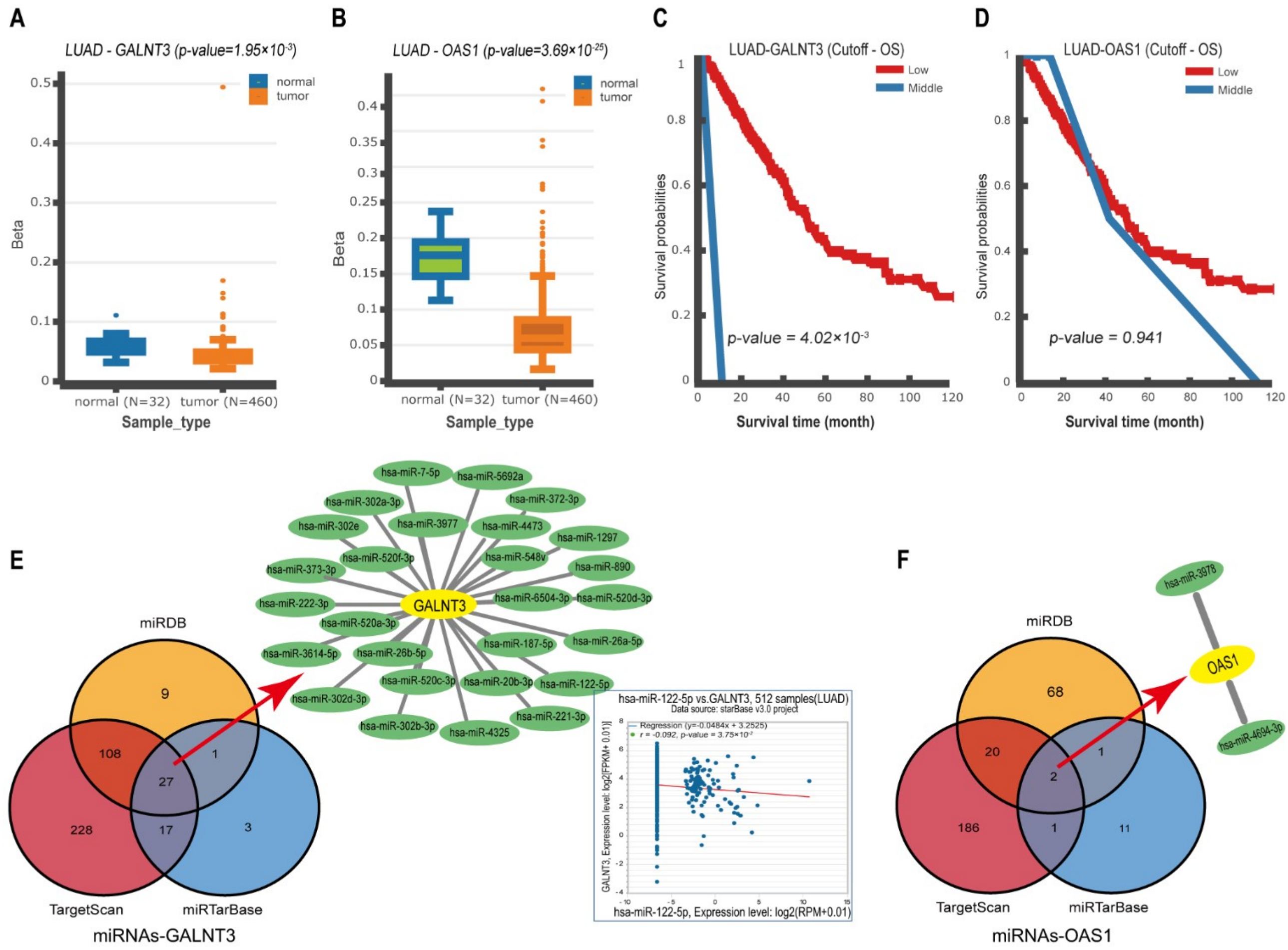

Fig. 5. DNA methylation status and miRNA regulation of *GALNT3* and *OAS1* in LUAD. (A,B) Boxplots showing the methylation status of the *GALNT3* and *OAS1* promoter regions in LUAD and in normal lung tissue. (C,D) Kaplan–Meier survival curves for LUAD patients according to the DNA methylation status of *GALNT3* and *OAS1*. (E) Potential miRNAs that target *GALNT3* mRNA were identified from the intersection of TargetScan, miRDB, and miRarBase databases. The correlation between *GALNT3* expression and hsa-miR-122-5p expression in LUAD was analyzed in the StarBase database. (F) Potential miRNAs that target *OAS1* mRNA were identified from the intersection of TargetScan, miRDB, and miRarBase databases.

quency), *MYO3B* (6%), *XIRP2* (21%), *ADAMTS12* (23%), *ZFHX4* (30%), *LRP1B* (32%), *USH2A* (35%), *MUC16* (40%), *TP53* (46%), and *ERCC5* (4%) (Fig. 7F). The top 10 co-occurring genes with OAS1 alteration were *CIT* (6%), *TMEM116* (2.6%), *CCDC63* (4%), *RPH3A* (5%), *TBX3* (5%), *KSR2* (5%), *MED13 L* (7%), *TBX5* (7%), *RBM19* (9%), and *LRRIQ1* (10%) (Fig. 7G).

GALNT3 and OAS1 mRNA expression was confirmed in LUAD and normal tissue from the TCGA database. They were again found to be significantly elevated in primary LUAD tissue compared with corresponding adjacent normal tissue (Fig. 7H,I). In addition, RT-PCR was used to examine *GALNT3* and *OAS1* mRNA expression levels in the normal cell line BEAS-2B and in the LUAD cell line H1975. This confirmed the overexpression of both genes in LUAD cells (Fig. 7J,K). The overexpression of *GALNT3* and *OAS1* in LUAD was further validated in the GEO and TCGA databases (Figs. 3E,H,7H,I).

ROC curve analysis showed AUC values of 0.9366 and 0.9759 for *GALNT3* and *OAS1* expression, respectively (Fig. 7L), thus highlighting their diagnostic value for LUAD. Logistic regression analysis for the combination of *GALNT3* and *OAS1* expression revealed an AUC of 0.9878 (Fig. 7L), indicating a very high diagnostic value for this biomarker combination in LUAD.

4. Discussion

LUAD is the most common type of lung cancer and shows multi-step development [37]. Despite continuous improvements in treatment, the prognosis of LUAD has not changed significantly [38]. The main reasons for the poor prognosis are inadequate early diagnosis and a lack of effective treatment targets. Since EMT is thought to be an essential process in the development of LUAD, EMT-related genes are likely to be an excellent candidate group for the discovery of novel biomarkers.

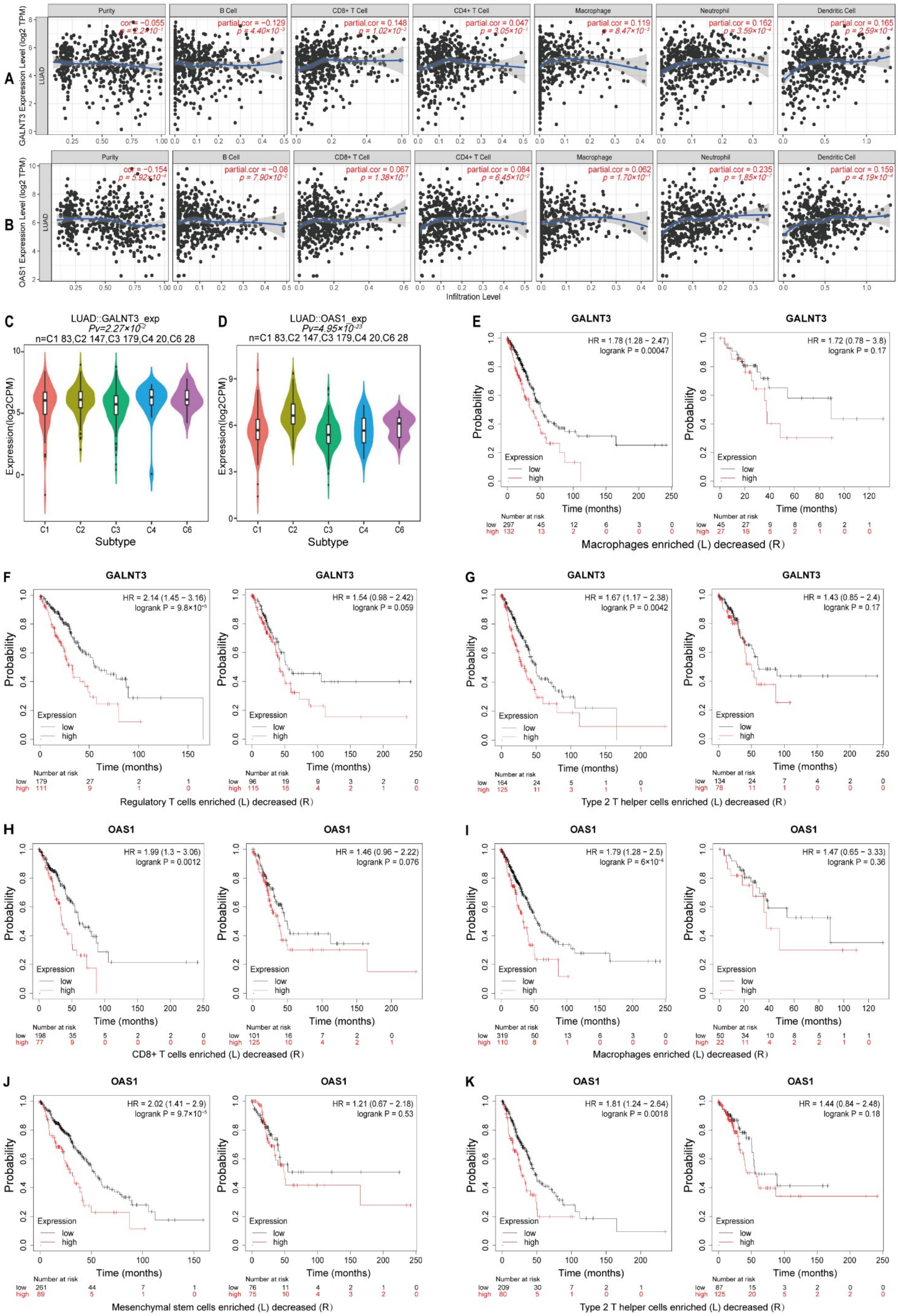

Fig. 6. Correlation analysis of *GALNT3* and *OAS1* expression in LUAD with infiltrating immune cells. (A,B) Correlation of *GALNT3* and *OAS1* expression with that of various immune cell markers. (C,D) Correlations between *GALNT3* and *OAS1* expression and immune subtypes in LUAD. (E–G) Kaplan–Meier survival analysis for LUAD patients with high or low *GALNT3* expression and according to the level of infiltration with different immune cell populations. (H–K) Kaplan–Meier survival analysis for LUAD patients with high or low *OAS1* expression and according to the level of infiltration with different immune cell populations. HR, hazard ratio.

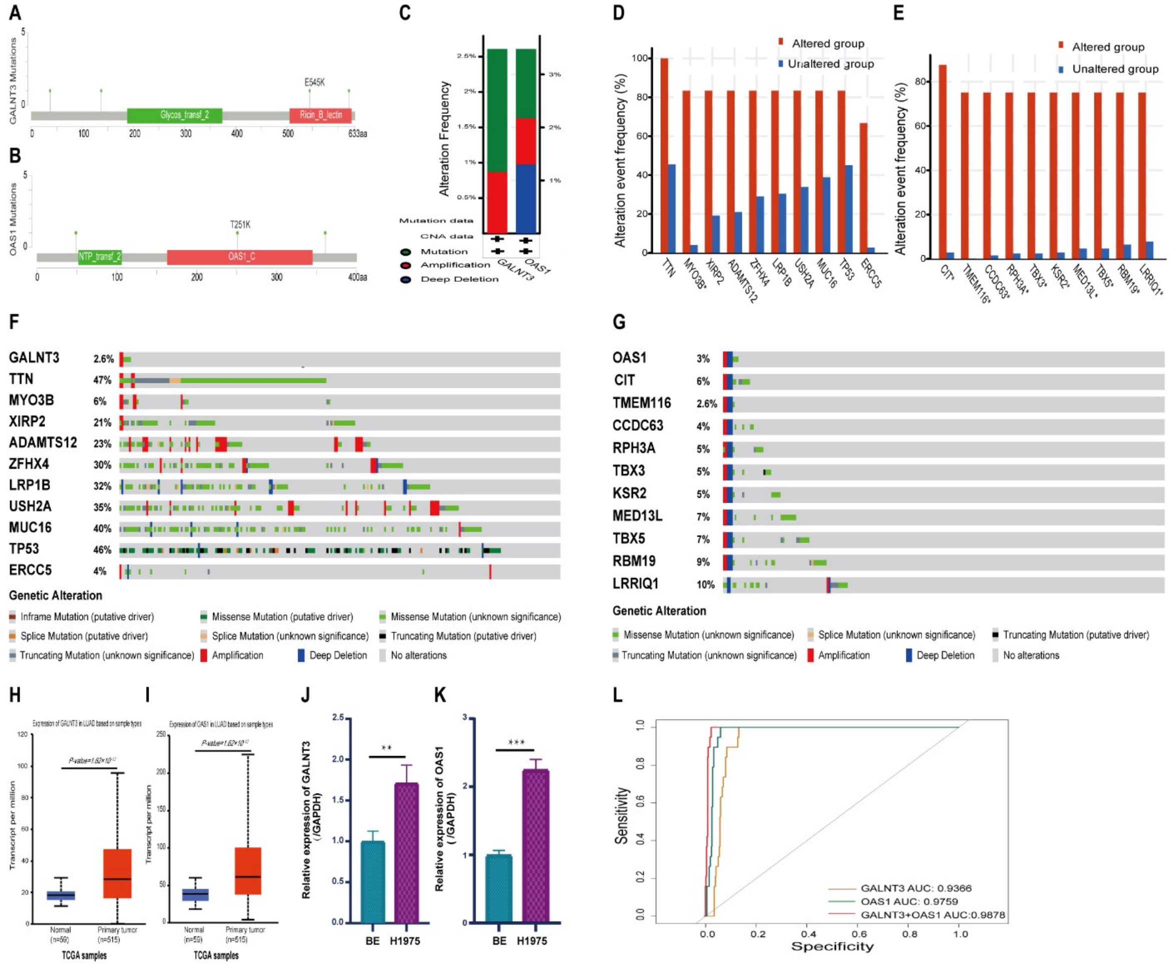

Fig. 7. DNA mutation, confirmation of mRNA expression, and diagnostic value of *GALNT3* and *OAS1* in LUAD. (A,B) Location of mutations in *GALNT3* and *OAS1* in LUAD. (C) *GALNT3* and *OAS1* genetic alteration types in LUAD. (D,E) Bar plots showing the top 10 gene alterations in LUAD with the highest frequency of co-occurrence with *GALNT3* and *OAS1* alterations. (F,G) Waterfall plots showing the co-occurrence pattern of *GALNT3* and *OAS1* gene mutations with the most frequent genetic changes in LUAD. (H,I) mRNA expression levels of *GALNT3* and *OAS1* in LUAD from the TCGA database. (J,K) mRNA expression levels of *GALNT3* and *OAS1* in LUAD cells (H1975) were higher than those in normal bronchial epithelial cells (BEAS-2B). (L) Receiver operating characteristic (ROC) curve analysis showing the diagnostic value of *GALNT3* and *OAS1* expression in LUAD, both individually and in combination (**$p<0.01$, ***$p<0.001$). BE, bronchial epithelial.

GALNT3, a member of the *GalNAc* transferase (*GALNAC-Ts*) gene family, is an EMT-related gene that is differentially expressed in various tumor types and has been implicated in cancer development [39–41]. This gene is highly expressed in ovarian and colon cancer [42–45]. *GALNT3* was more highly expressed in non-squamous cell carcinoma than in squamous cell carcinoma [46], which is consistent with the high expression of *GALNT3* in adenocarcinoma cell lines [47]. *OAS1* is a member of the 2′-5′ oligoadenylate synthetase (*OAS*) family, is also involved in EMT, and is highly expressed in bladder cancer and breast cancer [48–51]. In the present study, *GALNT3* and *OAS1* were found to be EMT-related genes through intersection of the GEO and EMTome databases. The expression of *GALNT3* and *OAS1* in LUAD was found to be positively correlated in the GEIPA database. We also found that mRNA and protein levels of *GALNT3* and *OAS1* in LUAD tissue were higher than in normal lung tissue. This was verified in H1975 cells using RT-PCR. Thus, we confirmed the differential expression of *GALNT3* and *OAS1* between LUAD and normal tissue, which should provide a solid foundation for future research. Univariate analysis showed that *GALNT3* and *OAS1* expression were both associated with poor prognosis in the GEO database, while multivariate analysis revealed that *GALNT3* expression, but not *OAS1*, was an independent prognostic marker in LUAD. We then comprehensively evaluated the prognostic significance of *OAS1* expression in LUAD by adding a multifaceted analysis method and database. Analysis of the UALCAN database revealed that *GALNT3* and *OAS1* expression were closely correlated with tumor stage, *TP53* mutation status, pathological N stage and patient survival. By comparing results from the GEO and UALCAN databases, we conclude the prognostic significance of *GALNT3* and *OAS1* expression may be related to multiple clinical factors, which could be valuable for further research.

EMT is one of the main mechanisms for the invasion and metastasis of cancer cells. It is an important biomarker for tumor prognosis and also one of the factors that affect response to treatment [52]. Previous research showed that overexpression of *GALNT3* was an independent prognostic factor for poor outcomes in patients with peripheral LUAD less than two centimeters in diameter [53]. *GALNT3* was also reported to be highly expressed in well-differentiated LUAD and associated with the degree of tumor differentiation, as well as being an independent factor of poor prognosis and early recurrence [54]. High *OAS1* expression in breast cancer and pancreatic cancer has been associated with a worse prognosis [49,55]. In the current study, high *GALNT3* and *OAS1* expression in LUAD was also associated with a worse prognosis, consistent with previous results in other cancer types. In addition, the expression of "response to drug and drug metabolism-other enzymes" genes was enriched in LUAD with high *GALNT3* and *OAS1* expression. These pathways are closely connected with resistance to anti-tumor drugs [56,57]. LUAD has a high frequency of *EGFR* mutations, while EGFR-TKI resistance remains a serious clinical problem [58]. Previous studies reported that T790M mutations and MET amplification were not present in tumor cells from EGFR-TKIs-resistant patients. Instead, these cells exhibit EMT features, with abnormal EMT-related molecular indicators and enhanced ability for migration and invasion [59,60]. Some researchers have also suggested that EMT may be a marker of resistance to EGFR-TKI [61,62]. EMT may therefore be a mechanism for acquiring resistance to EGFR-TKIs in lung cancer. The potential role of the EMT-related genes *GALNT3* and *OAS1* in EGFR-TKI resistance will thus be an interesting topic for future research. This may have practical value in the clinic, as well as provide a theoretical basis for overcoming EGFR-TKI resistance.

DNA methylation is the most widely studied epigenetic marker in cancer [63–65]. Deregulation of the normal methylation status of gene promoters can lead to aberrant gene expression [66,67]. In previous research, aberrant expression of *GALNT3* has been related to changes in DNA methylation [68]. In the present study, decreased methylation in the *GALNT3* promoter region was found to be associated with poor prognosis of LUAD patients. Promoter methylation of *OAS1* was inversely correlated with expression levels in LUAD, but not with poor prognosis. Genomic instability is also an important cause of tumorigenesis [69]. *GALNT3* gene amplification and mutations are found in LUAD and co-occur with mutations in essential genes such as *TTN*, *MYO3B*, *XIRP2*, *ADAMTS12* and *ZFHX4*. *OAS1* gene amplification, mutation and deletion were also observed, and these alterations co-occurred with mutations in *CIT*, *TMEM116*, *CCDC63*, *RPH3A* and *TBX3*. Therefore, alterations in the *GALNT3* and *OAS1*genes may affect the development and prognosis of LUAD. Upstream miRNAs can also affect the stability of target genes and negatively regulate the expression of downstream genes [70,71]. There is evidence that *GALNT3* and *OAS1* may be regulated by miRNA [72–74]. We identified 27 potential upstream miRNAs that target *GALNT3*, and two that target *OAS1*. Co-expression of hsa-miR-122-5p and *GALNT3* was observed, suggesting this upstream miRNA may regulate the expression of *GALNT3*. No co-expression was found between any miRNA and *OAS1* in the current research, although in future studies the miRNA search database should be expanded.

Infiltrating immune cells are a key part of the tumor microenvironment and have major effects on cancer development and prognosis [75,76]. GALNT3 and OAS1 expression are known to be closely associated with immune cell infiltration in various tumors [77–79]. In the present study, we confirmed that *GALNT3* and *OAS1* expression were associated with immune cells, immune cell markers, TILs, MHC molecules and immune modulators. High *GALNT3* expression was associated with poor prognosis in LUAD subgroups that were enriched with macrophages, regulatory T cells, and type 2 T helper cells. High *OAS1* expression was also associated with poor prognosis in LUAD that was enriched with CD8+ T cells, macrophages, mesenchymal stem cells, and type 2 T helper cells. These results suggest that *GALNT3* and *OAS1* expression may be associated with poor prognosis in LUAD because of their link to immune cell infiltration, however the specific mechanism requires further investigation.

There are several limitations to this study. First, the study samples were derived from public databases. Although strict screening criteria and repeated correction were applied, the quantity and quality of the tissue samples were limited. In-house follow-up data and increased sample size are needed to improve the quality of data in future studies. Second, this study only included LUAD samples, so the pathological type was restricted. Finally, this study explored the pathogenesis of *GALNT3* and *OAS1* using multiple omics only, and additional experimental studies are required to confirm the findings.

In summary, the EMT-related genes *GALNT3* and *OAS1* were highly expressed in LUAD and positively correlated. Their overexpression may be linked to gene mutation and methylation, and regulate the miRNA-*GALNT3/OAS1* network, which could in turn affect tumor infiltration with immune cells. The diagnostic value of *GALNT3* and *OAS1* expression combined was very high, and greater than that of the single markers. *GALNT3* and *OAS1* expression may therefore have clinical value for improving the accuracy of diagnosis and possibly also for improving the efficacy of treatment for LUAD.

5. Conclusions

The EMT-related genes *GALNT3* and *OAS1* are highly expressed in LUAD and are associated with significantly worse prognosis. The diagnostic value of the combined *GALNT3* and *OAS1* markers was greater than that of the individual genes. *GALNT3* and *OAS1* expressions in LUAD are likely to be affected by DNA mutations, upstream miRNA, and DNA methylation levels. These factors may affect the survival of LUAD patients by regulating immune cell infiltration into the tumor. Various aspects of *GALNT3* and *OAS1* expression were analyzed in this study, and these genes may be new molecular markers for the diagnosis and prognosis of LUAD. However, further studies are needed to confirm these findings.

Abbreviations

EMT, Epithelial-mesenchymal transition; NSCLC, Non-small-cell lung cancer; SCLC, Small cell lung cancer; LUAD, Lung adenocarcinoma; DEGs, Differentially expressed genes; *GALNT3*, Polypeptide N-Acetylgalactosaminyltransferase 3; *OAS1*, 2′-5′-Oligoadenylate Synthetase 1; TILs, Tumor infiltrates; MHC, Major histocompatibility class; CI, Confidence interval; HR, Hazard ratio; BP, Biological Process; CC, Cellular Component; MF, Molecular Function; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GEO, Gene Expression Omnibus; GEPIA, Gene Expression Profiling Interactive Analysis; TIMER, Tumor Immune Estimation Resource; HPA, Human Protein Atlas; UALCAN, University of Alabama at Birmingham Cancer; TCGA, The Cancer Genome Atlas; DNMIVD, DNA Methylation Interactive Visualization Database; TISIDB, Tumor–immune System Interactions Database; DAVID, Database for Annotation, Visualization and Integrated Discovery.

Availability of Data and Materials

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

Conceptualization, DL, LY, XF and QW; Data curation, DL, MF, MC, WW; Formal analysis, DL, LS, JY, XG; Funding acquisition, XF; Investigation, XG; Methodology, DL, JW, LY; Project administration, QW; Resources, WW; Software, DL, MF, MC, JY; Supervision, XF; Validation, DL, MF and YL; Visualization, DL; Writing—original draft, DL, JW; Writing—review and editing, LS, LY and QW. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work to take public responsibility for appropriate portions of the content and agreed to be accountable for all aspects of the work in ensuring that questions related to its accuracy or integrity.

Ethics Approval and Consent to Participate

Not applicable.

Acknowledgment

All the experiments were conducted at Laboratory Animal Center of Southwest Medical University and Model Animal and Human Disease Research of Luzhou Key Laboratory.

Funding

This study was funded by the Southwest Medical University School-level Scientific Research Project (No. 2021ZKQN095), the Science and Technology Development Fund, Macau SAR (File No.: 0098/2021/A2, 130/2017/A3, and 0099/2018/A3), the Science and Technology Planning Project of Guangdong Province (No. 2020B1212030008), the Natural Science Foundation of Sichuan Province (No.2022NSFSC0046), and Sichuan Science and Technology Program (No.2022YFS0631).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/j.fbl2810271.

References

[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA: a Cancer Journal for Clinicians. 2021; 71: 7–33.
[2] Marx A, Chan JKC, Coindre JM, Detterbeck F, Girard N, Harris NL, *et al*. The 2015 World Health Organization Classification of Tumors of the Thymus: Continuity and Changes. Journal of Thoracic Oncology: Official Publication of the International Association for the Study of Lung Cancer. 2015; 10: 1383–1395.
[3] Woodard GA, Jones KD, Jablons DM. Lung Cancer Staging and Prognosis. Cancer Treatment and Research. 2016; 170: 47–75.
[4] Nagy-Mignotte H, Guillem P, Vesin A, Toffart AC, Colonna M, Bonneterre V, *et al*. Primary lung adenocarcinoma: characteristics by smoking habit and sex. The European Respiratory Journal. 2011; 38: 1412–1419.
[5] Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. Annals of Global Health. 2019; 85: 8.
[6] Succony L, Rassl DM, Barker AP, McCaughan FM, Rintoul RC. Adenocarcinoma spectrum lesions of the lung: Detection, pathology and treatment strategies. Cancer Treatment Reviews. 2021; 99: 102237.
[7] Miller KD, Nogueira L, Mariotto AB, Rowland JH, Yabroff KR, Alfano CM, *et al*. Cancer treatment and survivorship statistics, 2019. CA: a Cancer Journal for Clinicians. 2019; 69: 363–385.
[8] Moore W, Talati R, Bhattacharji P, Bilfinger T. Five-year survival after cryoablation of stage I non-small cell lung cancer in

medically inoperable patients. Journal of Vascular and Interventional Radiology: JVIR. 2015; 26: 312–319.
[9] Burger GA, Danen EHJ, Beltman JB. Deciphering Epithelial-Mesenchymal Transition Regulatory Networks in Cancer through Computational Approaches. Frontiers in Oncology. 2017; 7: 162.
[10] Vander Ark A, Cao J, Li X. TGF-β receptors: In and beyond TGF-β signaling. Cellular Signalling. 2018; 52: 112–120.
[11] Yamaoka T, Kusumoto S, Ando K, Ohba M, Ohmori T. Receptor Tyrosine Kinase-Targeted Cancer Therapy. International Journal of Molecular Sciences. 2018; 19: 3491.
[12] Manfioletti G, Fedele M. Epithelial-Mesenchymal Transition (EMT) 2021. International Journal of Molecular Sciences. 2022; 23: 5848.
[13] Nieto MA, Huang RYJ, Jackson RA, Thiery JP. EMT: 2016. Cell. 2016; 166: 21–45.
[14] Brabletz S, Schuhwerk H, Brabletz T, Stemmler MP. Dynamic EMT: a multi-tool for tumor progression. The EMBO Journal. 2021; 40: e108647.
[15] Qu T, Zhang W, Qi L, Cao L, Liu C, Huang Q, *et al*. ISG15 induces ESRP1 to inhibit lung adenocarcinoma progression. Cell Death & Disease. 2020; 11: 511.
[16] Oh TI, Lee M, Lee YM, Kim GH, Lee D, You JS, *et al*. PGC1α Loss Promotes Lung Cancer Metastasis through Epithelial-Mesenchymal Transition. Cancers. 2021; 13: 1772.
[17] Sowa T, Menju T, Sonobe M, Nakanishi T, Shikuma K, Imamura N, *et al*. Association between epithelial-mesenchymal transition and cancer stemness and their effect on the prognosis of lung adenocarcinoma. Cancer Medicine. 2015; 4: 1853–1862.
[18] Pan G, Liu Y, Shang L, Zhou F, Yang S. EMT-associated microRNAs and their roles in cancer stemness and drug resistance. Cancer Communications (London, England). 2021; 41: 199–217.
[19] Tulchinsky E, Demidov O, Kriajevska M, Barlev NA, Imyanitov E. EMT: A mechanism for escape from EGFR-targeted therapy in lung cancer. Biochimica et Biophysica Acta. Reviews on Cancer. 2019; 1871: 29–39.
[20] Zhu X, Chen L, Liu L, Niu X. EMT-Mediated Acquired EGFR-TKI Resistance in NSCLC: Mechanisms and Strategies. Frontiers in Oncology. 2019; 9: 1044.
[21] Vasaikar SV, Deshmukh AP, den Hollander P, Addanki S, Kuburich NA, Kudaravalli S, *et al*. EMTome: a resource for pan-cancer analysis of epithelial-mesenchymal transition genes and signatures. British Journal of Cancer. 2021; 124: 259–269.
[22] Dennis G, Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, *et al*. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biology. 2003; 4: P3.
[23] Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Research. 2017; 45: W98–W102.
[24] Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. Trends in Genetics: TIG. 1997; 13: 163.
[25] Binder JX, Pletscher-Frankild S, Tsafou K, Stolte C, O'Donoghue SI, Schneider R, *et al*. COMPARTMENTS: unification and visualization of protein subcellular localization evidence. Database: the Journal of Biological Databases and Curation. 2014; 2014: bau012.
[26] Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, *et al*. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Research. 2017; 77: e108–e110.
[27] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, *et al*. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia (New York, N.Y.). 2017; 19: 649–658.
[28] Ding W, Chen J, Feng G, Chen G, Wu J, Guo Y, *et al*. DNMIVD: DNA methylation interactive visualization database. Nucleic Acids Research. 2020; 48: D856–D862.
[29] Agarwal V, Bell GW, Nam JW, Bartel DP. Predicting effective microRNA target sites in mammalian mRNAs. ELife. 2015; 4: e05005.
[30] Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Research. 2020; 48: D127–D131.
[31] Huang HY, Lin YCD, Li J, Huang KY, Shrestha S, Hong HC, *et al*. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Research. 2020; 48: D148–D154.
[32] Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Research. 2014; 42: D92–D97.
[33] Ru B, Wong CN, Tong Y, Zhong JY, Zhong SSW, Wu WC, *et al*. TISIDB: an integrated repository portal for tumor-immune system interactions. Bioinformatics (Oxford, England). 2019; 35: 4200–4202.
[34] Lánczky A, Győrffy B. Web-Based Survival Analysis Tool Tailored for Medical Research (KMplot): Development and Implementation. Journal of Medical Internet Research. 2021; 23: e27633.
[35] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, *et al*. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discovery. 2012; 2: 401–404.
[36] Barber RD, Harmer DW, Coleman RA, Clark BJ. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiological Genomics. 2005; 21: 389–395.
[37] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a Cancer Journal for Clinicians. 2018; 68: 394–424.
[38] Hsu CL, Chen KY, Shih JY, Ho CC, Yang CH, Yu CJ, *et al*. Advanced non-small cell lung cancer in patients aged 45 years or younger: outcomes and prognostic factors. BMC Cancer. 2012; 12: 241.
[39] Hang HC, Bertozzi CR. The chemistry and biology of mucin-type O-linked glycosylation. Bioorganic & Medicinal Chemistry. 2005; 13: 5021–5034.
[40] Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology. 2012; 22: 736–756.
[41] Magalhães A, Duarte HO, Reis CA. The role of O-glycosylation in human disease. Molecular Aspects of Medicine. 2021; 79: 100964.
[42] Wang ZQ, Bachvarova M, Morin C, Plante M, Gregoire J, Renaud MC, *et al*. Role of the polypeptide N-acetylgalactosaminyltransferase 3 in ovarian cancer progression: possible implications in abnormal mucin O-glycosylation. Oncotarget. 2014; 5: 544–560.
[43] Shibao K, Izumi H, Nakayama Y, Ohta R, Nagata N, Nomoto M, *et al*. Expression of UDP-N-acetyl-alpha-D-galactosamine-polypeptide galNAc N-acetylgalactosaminyl transferase-3 in relation to differentiation and prognosis in patients with colorectal carcinoma. Cancer. 2002; 94: 1939–1946.
[44] Sheta R, Bachvarova M, Plante M, Gregoire J, Renaud MC, Sebastianelli A, *et al*. Altered expression of different GalNAc transferases is associated with disease progression and poor prognosis in women with high-grade serous ovarian cancer. International Journal of Oncology. 2017; 51: 1887–1897.

[45] Liu B, Pan S, Xiao Y, Liu Q, Xu J, Jia L. LINC01296/miR-26a/GALNT3 axis contributes to colorectal cancer progression by regulating O-glycosylated MUC1 via PI3K/AKT pathway. Journal of Experimental & Clinical Cancer Research: CR. 2018; 37: 316.
[46] Dosaka-Akita H, Kinoshita I, Yamazaki K, Izumi H, Itoh T, Katoh H, *et al*. N-acetylgalactosaminyl transferase-3 is a potential new marker for non-small cell lung cancers. British Journal of Cancer. 2002; 87: 751–755.
[47] Nomoto M, Izumi H, Ise T, Kato K, Takano H, Nagatani G, *et al*. Structural basis for the regulation of UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyl transferase-3 gene expression in adenocarcinoma cells. Cancer Research. 1999; 59: 6214–6222.
[48] Gao L, Ren R, Shen J, Hou J, Ning J, Feng Y, *et al*. Values of OAS gene family in the expression signature, immune cell infiltration and prognosis of human bladder cancer. BMC Cancer. 2022; 22: 1016.
[49] Zhang Y, Yu C. Prognostic characterization of OAS1/OAS2/OAS3/OASL in breast cancer. BMC Cancer. 2020; 20: 575.
[50] Hovnanian A, Rebouillat D, Mattei MG, Levy ER, Marié I, Monaco AP, *et al*. The human 2',5'-oligoadenylate synthetase locus is composed of three distinct genes clustered on chromosome 12q24.2 encoding the 100-, 69-, and 40-kDa forms. Genomics. 1998; 52: 267–277.
[51] Kumar S, Mitnik C, Valente G, Floyd-Smith G. Expansion and molecular evolution of the interferon-induced 2'-5' oligoadenylate synthetase gene family. Molecular Biology and Evolution. 2000; 17: 738–750.
[52] Kong LR, Mohamed Salleh NAB, Ong RW, Tan TZ, Syn NL, Goh RM, *et al*. A common MET polymorphism harnesses HER2 signaling to drive aggressive squamous cell carcinoma. Nature Communications. 2020; 11: 1556.
[53] Zhao S, Guo T, Li J, Uramoto H, Guan H, Deng W, *et al*. Expression and prognostic value of GalNAc-T3 in patients with completely resected small (≤2 cm) peripheral lung adenocarcinoma after IASLC/ATS/ERS classification. OncoTargets and Therapy. 2015; 8: 3143–3152.
[54] Gu C, Oyama T, Osaki T, Li J, Takenoyama M, Izumi H, *et al*. Low expression of polypeptide GalNAc N-acetylgalactosaminyl transferase-3 in lung adenocarcinoma: impact on poor prognosis and early recurrence. British Journal of Cancer. 2004; 90: 436–442.
[55] Lu L, Wang H, Fang J, Zheng J, Liu B, Xia L, *et al*. Overexpression of OAS1 Is Correlated With Poor Prognosis in Pancreatic Cancer. Frontiers in Oncology. 2022; 12: 944194.
[56] Kaur G, Gupta SK, Singh P, Ali V, Kumar V, Verma M. Drug-metabolizing enzymes: role in drug resistance in cancer. Clinical & Translational Oncology: Official Publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico. 2020; 22: 1667–1680.
[57] Alfarouk KO, Stock CM, Taylor S, Walsh M, Muddathir AK, Verduzco D, *et al*. Resistance to cancer chemotherapy: failure in drug response from ADME to P-gp. Cancer Cell International. 2015; 15: 71.
[58] Passaro A, Jänne PA, Mok T, Peters S. Overcoming therapy resistance in EGFR-mutant lung cancer. Nature Cancer. 2021; 2: 377–391.
[59] Chung JH, Rho JK, Xu X, Lee JS, Yoon HI, Lee CT, *et al*. Clinical and molecular evidences of epithelial to mesenchymal transition in acquired resistance to EGFR-TKIs. Lung Cancer (Amsterdam, Netherlands). 2011; 73: 176–182.
[60] Weng CH, Chen LY, Lin YC, Shih JY, Lin YC, Tseng RY, *et al*. Epithelial-mesenchymal transition (EMT) beyond EGFR mutations per se is a common mechanism for acquired resistance to EGFR TKI. Oncogene. 2019; 38: 455–468.
[61] De Las Rivas J, Brozovic A, Izraely S, Casas-Pais A, Witz IP, Figueroa A. Cancer drug resistance induced by EMT: novel therapeutic strategies. Archives of Toxicology. 2021; 95: 2279–2297.
[62] Raoof S, Mulford IJ, Frisco-Cabanos H, Nangia V, Timonina D, Labrot E, *et al*. Targeting FGFR overcomes EMT-mediated resistance in EGFR mutant non-small cell lung cancer. Oncogene. 2019; 38: 6399–6413.
[63] Laird PW, Jaenisch R. DNA methylation and cancer. Human Molecular Genetics. 1994; 3 Spec No: 1487–1495.
[64] Mehta A, Dobersch S, Romero-Olmedo AJ, Barreto G. Epigenetics in lung cancer diagnosis and therapy. Cancer Metastasis Reviews. 2015; 34: 229–241.
[65] Papanicolau-Sengos A, Aldape K. DNA Methylation Profiling: An Emerging Paradigm for Cancer Diagnosis. Annual Review of Pathology. 2022; 17: 295–321.
[66] Robertson KD. DNA methylation, methyltransferases, and cancer. Oncogene. 2001; 20: 3139–3155.
[67] Klutstein M, Nejman D, Greenfield R, Cedar H. DNA Methylation in Cancer and Aging. Cancer Research. 2016; 76: 3446–3450.
[68] Vojta A, Samaržija I, Bočkor L, Zoldoš V. Glyco-genes change expression in cancer through aberrant methylation. Biochimica et Biophysica Acta. 2016; 1860: 1776–1785.
[69] Sieber O, Heinimann K, Tomlinson I. Genomic stability and tumorigenesis. Seminars in Cancer Biology. 2005; 15: 61–66.
[70] Shukla GC, Singh J, Barik S. MicroRNAs: Processing, Maturation, Target Recognition and Regulatory Functions. Molecular and Cellular Pharmacology. 2011; 3: 83–92.
[71] Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nature Reviews. Genetics. 2008; 9: 102–114.
[72] Lixin S, Wei S, Haibin S, Qingfu L, Tiemin P. miR-885-5p inhibits proliferation and metastasis by targeting IGF2BP1 and GALNT3 in human intrahepatic cholangiocarcinoma. Molecular Carcinogenesis. 2020; 59: 1371–1381.
[73] Che L, Yang H, Wang D, Liu S. Corylin sensitizes breast cancer cells to overcome tamoxifen resistance by regulating OAS1/miR-22-3p/SIRT1 axis. Acta Biochimica Polonica. 2021; 68: 757–764.
[74] Li Y, Zhao J, Zhang W, Wang A, Jiao M, Cai X, *et al*. LINC02535/miR-30a-5p/GALNT3 axis contributes to lung adenocarcinoma progression via the NF- κ B signaling pathway. Cell Cycle (Georgetown, Tex.). 2022; 21: 2455–2470.
[75] Zheng X, Hu Y, Yao C. The paradoxical role of tumor-infiltrating immune cells in lung cancer. Intractable & Rare Diseases Research. 2017; 6: 234–241.
[76] Liu X, Wu S, Yang Y, Zhao M, Zhu G, Hou Z. The prognostic landscape of tumor-infiltrating immune cell and immunomodulators in lung cancer. Biomedicine & Pharmacotherapy. 2017; 95: 55–61.
[77] Park MS, Yang AY, Lee JE, Kim SK, Roe JS, Park MS, *et al*. GALNT3 suppresses lung cancer by inhibiting myeloid-derived suppressor cell infiltration and angiogenesis in a TNFR and c-MET pathway-dependent manner. Cancer Letters. 2021; 521: 294–307.
[78] Jin K, Qiu S, Jin D, Zhou X, Zheng X, Li J, *et al*. Development of prognostic signature based on immune-related genes in muscle-invasive bladder cancer: bioinformatics analysis of TCGA database. Aging. 2021; 13: 1859–1871.
[79] Song C, Guo Z, Yu D, Wang Y, Wang Q, Dong Z, *et al*. A Prognostic Nomogram Combining Immune-Related Gene Signature and Clinical Factors Predicts Survival in Patients With Lung Adenocarcinoma. Frontiers in Oncology. 2020; 10: 1300.