Identification of a ZC3H12D-Regulated Competing Endogenous RNA Network for Prognosis of Lung Adenocarcinoma at Single-Cell Level

Wenhan Chen
Second Affiliated Hospital of Fujian Medical University
Zhifeng Guo
Second Affiliated Hospital of Fujian Medical University https://orcid.org/0000-0002-8715-2757
Jingyang Wu
Second Affiliated Hospital of Fujian Medical University
Guofu Lin
Second Affiliated Hospital of Fujian Medical University
Shaohua Chen
Second Affiliated Hospital of Fujian Medical University
Qinhui Lin
Second Affiliated Hospital of Fujian Medical University
Jiansheng Yang
Second Affiliated Hospital of Fujian Medical University
Yuan Xu
Second Affiliated Hospital of Fujian Medical University
Yiming Zeng (zeng_yiming@fjmu.edu.cn)
Second Affiliated Hospital of Fujian Medical University https://orcid.org/0000-0001-5635-3006

Primary research

Keywords: zinc finger CCCH-type containing 12D (ZC3H12D), competitive endogenous RNA, single-cell RNA sequencing, lung adenocarcinoma, immunomics.

Posted Date: June 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-603118/v1

Version of Record: A version of this preprint was published at BMC Cancer on January 28th, 2022. See the published version at https://doi.org/10.1186/s12885-021-08992-1.

Abstract

Objective To identify hub genes from the competing endogenous RNA (ceRNA) network of lung adenocarcinoma (LUAD) and to explore their potential function on prognosis of patients from a single-cell perspective.

Methods We performed RNA-sequencing of LUAD to construct ceRNA regulatory network, integrating with public databases to identify the vital pathways related to patients' prognosis and to reveal the expression level of hub genes under different conditions, the functional enrichment of co-expressed genes and their potential immune-related mechanisms.

Results ZC3H12D-hsa-miR-4443-ENST00000630242 axis was found to be related with LUAD. Lower ZC3H12D expression was significantly associated with shorter overall survival (OS) of patients (HR=2.007, P<0.05), and its expression was higher in early-stage patients, including T1 (P<0.05) and N0 (P<0.05). Additionally, ZC3H12D expression was higher in immune cells displayed by single-cell RNA-sequencing data, especially in Treg cells of lung cancer and CD8 T cells, B cells and CD4 T cells of LUAD. In the brain metastasized, the expression of ZC3H12D in macrophages was relatively abundant. The functional enrichment analysis showed that the co-expressed genes mainly played a role in lymphocyte activation and cytokine-cytokine receptor interaction. In addition, ZC3H12D was associated with multiple immune cells and immune molecules, including immune checkpoints CTLA4, CD96 and TIGIT.

Conclusion ZC3H12D-hsa-miR-4443-ENST00000630242 ceRNA network was identified in LUAD. ZC3H12D could affect the survival and prognosis of patients by regulating mRNA, miRNA, lncRNA, immune cells and immune molecules. Therefore, it may serve as a vital predictive marker and could be regarded as a potential therapeutic target for LUAD in the future.

Introduction

Lung adenocarcinoma (LUAD), accounting for approximately 60% of non-small cell lung cancer (NSCLC), is the most common subtype of lung cancer with a high incidence worldwide[1, 2]. With the advancement of technological innovation, the mechanisms behind LUAD are gradually revealed. Growing evidence highlights the vital role of the competitive endogenous RNA (ceRNA) regulatory networks in LUAD. For instance, AC079160.1-miR-539-5p-CENPF axis may participate in hypoxia-induced tumor cell stemness of LUAD. Low expression of AC079160.1 and CENPF and high expression of miR-539-5p were correlated with hypoxia and stemness index, indicating a better prognosis[3]. Furthermore, lnc01833-miR-519e-3p-S100A4 axis might promote LUAD progression, and lnc01833 overexpression can significantly improve proliferation and invasion ability of lung cancer cells as well as promote the EMT process[4].

Although the function of ceRNA network concerning LUAD has been uncovered gradually, the role of single-cell RNA sequencing (scRNA-seq) in LUAD is still waiting to be explored. ScRNA-seq is a novel hot spot to demonstrate the underlying mechanisms of LUAD, uncovering new differentially expressed genes as well as the heterogeneity of immune response-related genes[5, 6]. Besides, scRNA-seq was applied to

unravel the molecular and cellular reprogramming mechanisms in metastatic LUAD and the relationship between surfactant genes expression and overall survival in LUAD brain metastases patients[7, 8]. When it comes to immunotherapy, scRNA-seq technology might provide different expression aspect at single-cell level and identify heterogeneity in LUAD after being treated with anti-cancer drugs[9–11]. Therefore, integrating ceRNA network with scRNA-seq data may provide a promising strategy for further understanding the underlying mechanism of LUAD, and might collect more valuable knowledge to improve individual targeted treatment.

At present, there is no investigation on zinc finger CCCH-type containing 12D (ZC3H12D)-hsa-miR-4443-ENST00000630242 axis, which might play a critical role in LUAD. Anti-oncogene ZC3H12D, also called p34, is an RNase of the Regnase-1 family that is associated with gene expression such as IER3, TNF, IL-6, NF-κB and TLR, degrading inflammatory transcripts and attenuating macrophage response[12–14]. Previous studies have demonstrated that ZC3H12D could be targeted by miR-128-3p, involving in cell proliferation and migration in osteosarcoma [15]. However, the functions of hsa-miR-4443 in different cancers are heterogeneous. It could promote the drug resistance of NSCLC to epirubicin[16]. While, the overexpression of hsa-miR-4443 acts in a tumor-suppressive manner, decreasing the invasiveness of hepatocellular carcinoma (HCC)[17], glioblastoma (GBM)[18] and colorectal cancer (CRC)[19]. Next, long non-coding RNAs (lncRNAs) have demonstrated their roles in ceRNA networks as regulators and participated in the regulation of various pathological processes related to cancers. Compared with mRNA and miRNA, little is known about the function of lncRNA FAM30A. Highly expressed in B cells, FAM30A is correlated with the regulation of the immune response and immunoglobulin genes[20, 21]. Our study aims to reveal the function of the ZC3H12D-hsa-miR-4443-ENST00000630242 axis in the prognosis of LUAD, especially the role of ZC3H12D concerning immunomics.

METHODS

Patients And Clinical Samples

Patients and clinical samples

10 paired LUAD and paracancerous tissues were collected between January, 2019 and May, 2019 at Fujian Medical University Second Affiliated Hospital. All specimens were obtained from LUAD patients who only received primary surgical treatment. These specimens were immediately snap-frozen by liquid nitrogen after resection, and then stored at -80℃ until RNA extraction. The clinicopathological diagnosis was confirmed by two well-experienced pathologists according to the World Health Organization (WHO) guidelines (2015). The study was approved by the bioethical committees at The Second Affiliated Hospital of Fujian Medical University, China (2020-206). And, all participating patients provided written informed consent.

RNA extraction and sequencing

Total RNA of LUAD tissues and paracancerous tissues was isolated from collected frozen tissues with the RNeasy Mini Kit (Qiagen, Germany) following the standard manufacturer's instructions. Qubit 4.0 (Thermo Fisher Scientific, Wilmington, DE, USA) was used to evaluate the RNA concentration, and agarose gel electrophoresis was applied to assess the RNA quality.

Then, ribosomal RNA was removed from the total RNA to retain the maximum residual non-coding RNA (ncRNA). After fragments of rRNA-depleted RNA, the cDNA library construction was performed with the TruSeq RNA Sample Prep Kit (Illumina, San Diego, CA, USA). LncRNA/mRNA sequencing libraries were generated using the VAHTS total RNA-seq Library Prep kit for Illumina (Vazyme NR603, China) following the standard manufacturer's instructions. The cDNA fragments with 150-bp paired-end reads were generated for RNA sequencing. Then, NEBNext® Multiplex Small RNA Library Prep Set for Illumina® (NEB) was applied to establish the miRNA library for samples. 12 libraries were pooled and sequenced in a single lane of Illumina HiSeq Xten sequencing platform. And the Illumina's TruSeq small RNA library preparation kit was applied to establish the miRNA library with 50-bp paired-end reads generated. After library construction, the sequencing for both lncRNA/mRNA and miRNA was carried out with the Illumina HiSeq Xten platform.

Identifications of differentially expressed mRNAs, miRNAs and lncRNAs

Mirdeep2 (v2.0.0.5) was applied to predict new miRNA, whose expression was calculated and standardized using counts per million (CPM) read. While, LncRNAs were annotated by three databases, including CNCI (https://github. com/www-bioinfo-org/CNCI) and CPC2 (http://cpc2.cbi. pku.edu.cn/), CPAT (https://sourceforge.net/projects/rna-cpat/), PLEX (https://sourceforge.net/ projects/plek/). The intersection was retained for further analysis.

DESeq2Rpackage(https://bioconductor.org/packages/release/bioc/html/DESeq2.html) in the Bioconductor project was applied to screen the differentially expressed mRNAs (DEmRNAs), differentially expressed miRNAs (DEmiRNAs) and differentially expressed lncRNAs (DElncRNAs) between LUAD and normal tissues. |Log2(fold change)|(|log2FC|) ≥1 and statistical P<0.05 were set as cut-off criteria. Subsequently, unsupervised hierarchical clustering was performed for DE-RNAs, and the expression patterns of which in paired tissues were displayed in form of heatmap using the 'pheatmap' R package (https://cran.r-project.org/ web/ packages/pheatmap/index.html).

Analysis of the DElncRNAs Enrichment pathway

Gene ontology (GO, http://www.bioconductor.org/packages/release/bioc/ html/topGO.html) function analysis was performed to screen enrichment of targeted genes to annotate the biological functions regulated by DElncRNAs, including biological processes (BP), cellular component (CC), and molecular function (MF). In addition, Kyoto Encyclopedia of Genes and Genomes (KEGG,

http://www.genome.jp/keggbin/show_organism?menu_type=pathway_maps&org=hsa) analysis was conducted to determine the vital signaling pathways associated with DElncRNAs. Both GO and KEGG enrichment analysis set gene count ≥2 and P value <0.05 as the threshold for statistical significance.

Predication of miRNA Regulation Relationship

The miRWalk 2.0 (http://zmf.umm.uni-heidelberg.de/ pps/zmf/mirwalk2/) was applied to perform the prediction of miRNA-gene analysis of DEmiRNA. Additionally, miRWalk, miRanda, miRDB, miRMap and TargetScan databases were applied to predict the potential DEmiRNA-DEmRNA regulatory relationships. Subsequently, the StarBase (http://starbase.sysu.edu.cn/) database was used to predict the potential miRNA-lncRNA regulatory relationships by DEmiRNA. Then, the DEmRNA-DEmiRNA and DEmiRNA-DElncRNA regulatory relationships were successfully constructed based on shared DEmiRNAs with which DEmRNAs and DElncRNAs interact, and were visualized by Cytoscape software.

Construction of lncRNA-miRNA-mRNA ceRNA network

Based on the hypothesis of miRNA sponge, the positive correlation expression of DElncRNAs-DEmRNAs was focused on, and the positive correlation co-expressed relationships between DEmRNAs and DElncRNAs simultaneously regulated by DEmiRNAs was obtained. Subsequently, the competing endogenous RNA (CeRNA) network was constructed based on shared miRNAs. Furthermore, we applied a hypergeometric cumulative distribution function test to predict the possible ceRNA pairs, and only the pairs with correlation coefficient□0.5 and P value□0.05 were selected. The Sanguini diagram was built based on the R software package ‘ggalluval’ (https://github.com/corybrunson/ggalluvial).

Public data source

Raw counts of RNA-sequencing data from 510 LUAD and paired normal tissue samples and corresponding clinical information were obtained from The Cancer Genome Atlas (TCGA) dataset (https://portal.gdc.cancer.gov/) in January 2020, and the method of acquisition and application for data were complied with the guidelines and policies. A total of 515 paired LUAD tissues RNA-seq data in the TCGA database were retrieved. LUAD-related RNA-sequencing data and corresponding clinical information were obtained from the TCGA database as following criteria: 1) histologically diagnosed as LUAD; 2) data with clinical information. Finally, a total of 510 paired LUAD tissues was included for further analysis.

MiRNAs data was downloaded from KM-plotter (https://kmplot.com/ analysis/)[22], and LncRNAs data was obtained from lnCAR database(https:// lncar.renlab.org/explorer)[23]. The single-cell sequencing data of normal lung tissue from 8 mice and 3 humans came from Tabula Muris (https://tabula-muris.ds.czbiohub.org/)[24] and synapse database(https://www.synapse.org/), respectively. While the

single-cell sequencing data of NSCLC, including LUAD, resulted from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) and European Molecular Biology Laboratory (EMBL, https://www.ebi.ac.uk/). The databases were included only if they contained single cell RNA sequencing of NSCLC. Finally, we founded 6 valuable datasets. The ScRNA-seq data of EMTAB6149 was prepared from 5 dissected lung tumors[25], while GSE117570 contained scRNA-seq data of tumor-infiltrating immune cells in the early-stage of NSCLC[26]. Besides, GSE127465 contained scRNA-seq data of red blood cell (RBC)-depleted cells from NSCLC tumor and from blood of 7 patients, as well as CD45-positive cells from lungs of 2 healthy mice and 2 lung tumor-bearing mice[27]. GSE127471 data were collected from cryopreserved peripheral blood mononuclear cells of NSCLC and GSE131907 contained data of 208,506 cells derived from 44 patients with LUAD, taken from normal lung tissues, primary tumor, normal lymph nodes, invaded lymph nodes, pleural effusion, and brain metastases[28, 29]. Additionally, GSE139555 contained data of pretreatment samples from 14 NSCLC patients, which covered normal tissue, primary tumor and peripheral blood[30]. The present study meets the criteria of data usage and publishing of the National Cancer Institute of National Institutes of Health, and no approval from the ethics committee was required.

Survival prognosis and clinical factors

The Kaplan–Meier survival analysis with log-rank test was used to compare the survival difference between the above two groups using ‘survival’ (https:// cran.r-project.org/web/packages/survival/index.html) and ‘survminer’ (https:// cran.r-project.org/web/packages/survminer/index.html) R packages. TimeROC analysis was performed to compare the predictive accuracy of each gene and risk score. For Kaplan–Meier curves, p-values and hazard ratio (HR) with 95% confidence interval (CI) were generated by log-rank tests. Univariate and multivariate cox regression analysis were performed to identify the proper terms to build the nomogram. The forest plot was used to display *P* values, HR and 95% CI of each variable using ‘forestplot’ R package. A nomogram was developed based on the results of multivariate Cox proportional hazards analysis to predict the X-year overall survival (OS). The nomogram provided a graphical representation of the factors, which can be used to calculate the risk of recurrence for an individual by the points associated with each risk factor through ‘rms’ R package (https://cran.r-project.org/web/packages/rms/index.html). All the above analysis methods and R package were implemented by R foundation for statistical computing (2020) version 4.0.3 and ggplot2 (v3.3.2). *P* value<0.05 was considered statistically significant.

Gene expression atlas based on single-cell sequencing data

We chose the t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm to reduce the dimensionality of the quality-controlled scRNA-seq data of normal lung cell from mouse. Moreover, scRNA-seq data from lung cell atlas and lung cancer brain metastases of human were re-analyzed through the UCSC cell browser (https://cells.ucsc.edu/)[31].To understand the expression of ZC3H12D in different cell types

across selected datasets, we applied TISCH (http://tisch.comp-genomics.org/) to obtain the ZC3H12D average expression data from multiple databases[32]. Furthermore, we explored the expression of ZC3H12D in LUAD and other well-characterized NSCLC at single-cell level and identified the distribution of expression of ZC3H12D in different crucial cell-types across datasets.

Screening of co-expressed genes and enrichment analysis

We applied Spearman's correlation analysis to identify genes that were related to the expression level of ZC3H12D. Those genes with correlation coefficient 0.5, *P*< 0.01 and FDR < 0.01 were considered as co-expressed genes for GO functional enrichment analysis and KEGG pathway enrichment analysis. The top 20 categories by GO analysis and top 20 pathways enriched by KEGG analysis were displayed, respectively. The volcano plot and heat map were drawn by LinkedOmics (http://www.linkedomics.org/login.php)[33]and Metascape (http://metascape.org/)[34], and the bubble diagram was drawn by 'ggplot2' R package(https://cran.r-project.org/web/packages/ggplot2/index.html).

The immune cell infiltration level and immune molecules expression

To obtain reliable immune infiltration estimations, we utilized the 'immunedeconv'(https://www.rdocumentation.org/packages/immunedeconv/versions/2.0.0) R package that integrated six state-of-the-art algorithms, including Tumor Immune Estimation Resource (TIMER, http://timer.cistrome.org/)[35]. Besides, immunostimulators and immunoinhibitors were selected and the expression levels (transcripts per kilobase of exonmodel per million mapped reads) of these genes were extracted. The two-gene, and the multi-gene correlation map was displayed using 'ggstatsplot' (https://cran.r-project.org/ web/ packages/ggstatsplot/index.html) and 'pheatmap' R package, respectively. Additionally, we used Spearman's correlation analysis to describe the correlations among gene expression. And *P*<0.05 was considered statistically significant.

Results

Construction of ceRNA regulatory network in LUAD

Totally, 49 mRNAs, 99 miRNAs and 50 lncRNA were obtained to construct ceRNA regulatory networks (Figure 1a). To identify a powerful pathway from this complicated ceRNA regulatory network, we focused on the function of lncRNA-miRNA-mRNA axises that could predict OS of LUAD. Verified by the Kaplan–Meier survival analysis, five mRNAs including ZC3H12D were identified to be significant associated with OS, including ZC3H12D, GNAO1, KSR2, SBK1 and SLIT3. (Figure 1b and Supplement 1). Compared with TCGA normal data, most tumor samples had higher expression of SBK1 and ZC3H12D, while lower SLIT3 expression was observed in LUAD. However, there was no significant difference in the expression of GNAO1 and KSR2 between tumor and paracancerous tissue samples. (Supplement 1). Meanwhile, both SBK1 and SLIT3 were found to play a vital role in LUAD by miRNA or lncRNA. Thereinto, we focused on ZC3H12D to construct a ceRNA regulatory network, and screened for potentially functional miRNAs and

lncRNAs (Figure 1c). Then, high expression of hsa-miR-4443 was not conducive to the prognosis of patients with LUAD (Figure 1d), but high expression of ENST00000630242 was beneficial to the prognosis of patients with LUAD (Figure 1e). Furthermore, ZC3H12D-hsa-miR-4443-ENST00000630242 axis was established, which might play a crucial role in the prognosis of LUAD patients.

Establishment of Cox prognostic model

To explore the predictive role of ZC3H12D in LUAD patients, we constructed Cox regression models. Fortunately, ZC3H12D was beneficial to the overall survival of patients in both univariate and multivariate Cox regression models. (Figure 2a and Figure 2b). Based on the identifications from the multivariate cox proportional hazards analysis, a prediction nomogram was developed to calculate the risk of recurrence for an individual by the points associated with the three risk factors (ZC3H12D, N stage, T stage). The total score ranged from 0 to160, and was calculated by summing all the scores in each variable. According to the score calculated, we could roughly predict one-year survival rate, three-year survival rate and five-year survival rate of patients. A higher score indicated a higher risk of death (Figure 2c). For the purpose of internal validation, the calibration plot showed that the nomogram predicting 1-year OS was relatively more accurate than the nomogram predicting 1-year OS,2-year OS and 3-year OS of LUAD patients (Figure 2d).

Relationship between vital clinical factors and ZC3H12D expression level

As identified from the prognostic model, ZC3H12D, N stage and T stage could predict OS of LUAD patients. To explore the relationship between ZC3H12D expression and clinical factors associated with prognosis, we analyzed the significance *P* values by chi-square test. The results showed that T and N stages displayed significant differences between in group C1 and in group C2. In group C1, the expression of ZC3H12D was significantly higher than the median, while in group C2 was significantly lower than the median. (Figure 3a and Figure 3b). From the correspondent bar chart, it exhibited that the proportion of ZC3H12D high expression in T1 and N0 was significantly higher than that of ZC3H12D low expression (Figure 3c and Figure 3d). To further verify the results, we combined N1, N2 and N3 group into one group, which was compared with N0 group to observe the difference of expression between the two groups. In N0 group, the ZC3H12D expression was significantly higher than another group (Wilcoxon rank sum test, *P*=8.3e-05) (Figure 3e). In addition, eliminating unknown samples, we took T2, T3 and T4 as a whole. Compared with T1, we found that there was a statistical significance between them and ZC3H12D expression was relatively high in T1 phase (Figure 3f).

ScRNA sequencing reveals ZC3H12D expression in normal lung tissue and LUAD

Different from conventional bulk RNA-sequencing, we applied scRNA-seq database to explore the expression level of ZC3H12D in normal lung tissue and LUAD using. We found that ZC3H12D was mainly expressed in immune cells, which can be found in both human and mouse lung specimens (Figure 4a-c

and Supplement 2). Using scRNA-seq data from the GEO database, we found that ZC3H12D expression displayed heterogeneity in different clusters of cells in different NSCLC datasets (Figure 4d). ZC3H12D was more abundantly expressed in CD4 T cells using GSE127465 and in Treg cells using GSE99254 (Figure 4e-h). Then, to eliminate the effect of lung squamous cell carcinoma we used another scRNA-seq dataset from GSE131907 to unravel the expression level in LUAD, indicating that it was relatively highly expressed in CD8 T cells, B cells and CD4 T cells (Figure 4i and Figure 4j). Also, it was expressed in some plasma cells, DC cells, monocytes or macrophages, while ZC3H12D was not abundantly expressed in fibroblasts, mast and endothelial cells. However, when lung cancer metastasized to the brain, the expression levels of various types of cells changed. The ZC3H12D expression in macrophages was significantly increased, even higher than that in CD4 and CD8 T cells (Figure 4k).

Screening for important co-expressed genes

Base on Spearman correlation test, we identified 19,987 genes from the TCGA LUAD data. A total of 12,201 genes were positively correlated with ZC3H12D expression, while 7,786 genes were negatively correlated with ZC3H12D expression, which displayed in the volcanic map (Figure 5a). The top 50 positively correlated significant genes and the top 50 negatively correlated significant genes were screened and plotted as heat maps, respectively (Figure 5b and Figure 5c).

GO functional enrichment analysis and KEGG pathway enrichment analysis

After further filtration of 12,201 positive related genes, 345 genes were obtained using correlation coefficient > 0.5, p-value < 0.01 and FDR < 0.01 as criteria for screening Gene ontology (GO) classification functional enrichment and Kyoto encyclopedia of genes and genomes (KEGG) classification pathway enrichment of 345 genes were performed. The results demonstrated that these genes were mainly enriched in the immune function of lymphocyte activation, antigen receptor-mediated signaling pathway, B cell activation a and alpha-beta T cell activation (Figure 5d and Supplement 3). Meanwhile, co-expressed genes also participated in many pathways, such as cytokine-cytokine receptor interaction, cell adhesion molecules and Th17 cell differentiation (Figure 5e).

Correlation between gene expression and immune cell infiltration

In the face of so much evidences showing the potential relationship between ZC3H12D and immunity, we decided to explore the correlation between expression of ZC3H12D and various immune cell infiltration level in LUAD. Based on the experimental peritoneal cancer index (EPCI) algorithm, six major immune cells showed a trend: the score of the G1 group with high expression was higher than the score of the G2 group with low gene expression, interestingly, in the undefined cells, the score of the G1 group with high gene expression was relatively lower than that of the G2 group with low gene expression (Figure 6a).

Regulation between ZC3H12D and important immune molecules

Among the co-expressed genes screened, ZC3H12D expression is linked to many immune-related genes. To further explore its regulatory effect, we analyzed the correlation between the expression of ZC3H12D and the common immunostimulators. Ultimately, we identified 10 immunostimulators (CD27, CD28, CD40LG, CD48, CXCR4, ICOS, KLRK1, LTA, TNFRSF13B, TNFRSF13C) with statistical significance (Figure 6b). Interestingly, we also found that the expression of ZC3H12D was related to immunoinhibitors. 10 statistically significant immunoinhibitors were also screened, including several immune-checkpoint–relevant transcripts, such as CTLA4, CD96, TIGIT. However, the correlation between the immune checkpoint SIGLEC15 and ZC3H12D expression is relatively weak (Figure 6c).

Discussion

Screened from the ceRNA regulatory network of LUAD, ZC3H12D-hsa-miR-4443-ENST00000630242 pathway was found, which was closely related to the survival of patients. High expression of hsa-miR-4443 was not conducive to the prognosis of patients with LUAD, but high expression of ENST00000630242 was beneficial to the prognosis of patients with LUAD. Besides, high ZC3H12D expression was linked to various clinical characteristics and better prognosis. Enriched in immune cells, ZC3H12D was associated with various immune cell infiltration levels and immune molecules. The functional enrichment analysis also showed that the co-expressed genes mainly played a role in lymphocyte activation and cytokine-cytokine receptor interaction.

ZC3H12D is also called MCPIP4, C6orf95, TFL and dJ281H8.1. It is a tumor suppressor gene, which plays a critical role in many cancers, including tongue cancer[36], osteosarcoma[37] and lung cancer. As for lung cancer, most of previous studies focused on the effect of genetic polymorphisms of ZC3H12D, and little was known about its potential regulatory mechanisms on lung cancer[38]. Previous studies have demonstrated that ZC3H12D was associated with memory T lymphocytes and macrophages, participating in the regulation of inflammation[39, 51]. Therefore, we hypothesized that this gene might have an effect on LUAD through immune regulatory mechanism. In our study, we found that many immune-related genes were positively correlated with the expression of ZC3H12D, such as ZNF831, SLAMF1 and IL-16. Both ZNF831 and ZC3H12D were linked to Zinc Finger Family, and it has been reported that ZNF831 was specifically significant in the high immunity subtype of triple-negative breast cancer, which was characterized by anti-tumor immune activities, better immune cell infiltration and greater probability of OS[41]. And SLAMF1 could both inhibit proliferation and impair responses to B cell receptor ligation in IGHV mutated chronic lymphocytic leukemia, which was similar to the anti-tumor effect of ZC3H12D[42]. In the previous literature, ZC3H12D could regulate IL-6, which was a member of the interleukin family, and we further found that IL-16, another member of the interleukin family, is also closely related to ZC3H12D[43, 44].As a pro-inflammatory cytokine, IL-16 was associated with high grade immune related adverse events in advanced NSCLC treated with immune checkpoint inhibitors [45].

In addition to the correlation of co-expressed genes, it is more convincing to analyze the function and pathway of positively related gene sets in LUAD samples. In our study, a lot of genes participated in the functions and pathways associated with immunity. The function enrichment results showed that many

co-expressed genes were involved in lymphocyte activation and antigen receptor-mediated signaling pathway, which were consistent with the results of previous studies[46, 47]. In addition to the immune-related pathways, we also found that the co-expressed genes of ZC3H12D expression were enriched in well-known cancer-related pathways, such as NF-kB signaling pathway, Ras signaling pathway and Jak-STAT signaling pathway, which need to validate its function in cancer in further study.

Besides, the scRNA-seq technology provided us an innovative method to reveal the gene expression level in immune cells under different conditions[48, 49]. Based on the ScRNA-seq in our study, it showed that ZC3H12D was not homogeneous among different clusters in tumor, but it selectively highly expressed in immune cells. Compared with the samples of normal lung tissue, LUAD tissue, non-small cell lung cancer and brain metastasis of lung cancer, we found that the expression of ZC3H12D in different immune cell types, including conventional CD4 T cells, regulatory T cells, monocytes or macrophages, would change under different conditions, which reflected the plasticity of to a certain extent. Traditional bulk RNA-sequencing also showed that when it was highly expressed, the EPIC scores of the main immune cells were correspondingly higher. Meanwhile, the high expression of ZC3H12D was also common in T1 and N0, and was related to some immune molecules, so we speculated that it may play an anti-cancer effect by regulating immunity in the early stage of LUAD. Future studies are required to compare the immune changes caused by ZC3H12D between early-stage and advanced-stage, so as to further reveal the function of ZC3H12D on the dynamic heterogeneity of LUAD.

Apart from immune mechanism, ceRNA regulatory network may also involve in the prognosis of LUAD. In the present study, lncRNA ENST00000630242, acting as a ceRNA, could “sponge” hsa-miR-4443 to regulate the expression of target ZC3H12D. Overexpression of ZC3H12D was beneficial to prolong the survival time of LUAD patients, while hsa-miR-4443 was not conducive to the prognosis of patients, which were conformity with the pertinent literature[16, 50]. Although there were few studies on ENST00000630242, the data showed that it was beneficial to the prognosis of patients with LUAD. Therefore, ZC3H12D-hsa-miR-4443-ENST00000630242 axis could be served as a novel potential target for LUAD treatment.

The major limitation of this study is that we have not confirmed the ENST00000630242-hsa-miR-4443-ZC3H12D axis by experiments, though the correlations have been primarily uncovered through RNA-seq from clinical samples. Although the function of ZC3H12D was revealed in the present study, the mechanisms of ENST00000630242 and hsa-miR-4443 in LUAD are still unclear. Therefore, further study is needed to explore the role of ENST00000630242 and hsa-miR-4443 in LUAD.

In summary, we found that ENST00000630242-hsa-miR-4443- ZC3H12D axis might be involved in the OS of LUAD patients. ZC3H12D, the core part of this pathway, was combined with some clinical factors to establish a cox model together. Furthermore, ZC3H12D expression at single level was unraveled in both normal lung tissues and lung tumors. Also, we found ZC3H12D expression was associated with some clinical features, important functions and pathways. Meanwhile, we explored the correlation between ZC3H12D and immune mechanisms to understand LUAD better.

Declarations

Ethics approval and consent to participate: The study was approved by the bioethical committees at The Second Affiliated Hospital of Fujian Medical University, China (2020-206). And the present study meets the criteria of data usage and publishing of the National Cancer Institute of National Institutes of Health, and no approval from the ethics committee was required.

Consent for publication All participating patients provided written informed consent.

Availability of data and materials The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests The authors declare that they have no competing interests" in this section.

Funding We acknowledge funding from funding project of Fujian Medical University College Student Innovation and entrepreneurship training program (Grant No. C19046) and funding project of Fujian Medical University College Student Innovation and entrepreneurship training program (Grant No. C19071).

Authors' contributions WC, ZG and JW were the major contributors in writing the manuscript. All authors read and approved the final manuscript."

Acknowledgments Not applicable

References

1. Bray F, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. Cancer J Clin. 2018;68(6):394–424.
2. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA: A Cancer Journal for Clinicians, 2019. 69(1): p. 7–34.
3. Guo L, et al. Construction and investigation of a combined hypoxia and stemness index lncRNA-associated ceRNA regulatory network in lung adenocarcinoma. BMC Med Genom. 2020;13(1):166–6.
4. Zhang Y, et al., The Long Noncoding RNA Linc01833 Enhances Lung Adenocarcinoma Progression via MiR-519e-3p/S100A4 Axis. Cancer management and research, 2020.
5. Zeng H, et al. Stemness Related Genes Revealed by Network Analysis Associated With Tumor Immune Microenvironment and the Clinical Outcome in Lung Adenocarcinoma. Frontiers in genetics. 2020;11:549213.
6. He D, et al. Single-cell RNA sequencing reveals heterogeneous tumor and immune cell populations in early-stage lung adenocarcinomas harboring EGFR mutations. Oncogene. 2021;40(2):355–68.
7. Liu Y, et al. Single-cell transcriptome analysis demonstrates inter-patient and intra-tumor heterogeneity in primary and metastatic lung adenocarcinoma. Aging. 2020;12(21):21559–81.

8. Ruan H, et al. Circulating tumor cell characterization of lung cancer brain metastases in the cerebrospinal fluid through single-cell transcriptome analysis. Clinical translational medicine. 2020;10(8):e246.
9. Guo L, et al. TOX correlates with prognosis, immune infiltration, and T cells exhaustion in lung adenocarcinoma. Cancer medicine. 2020;9(18):6694–709.
10. Cai XC, et al. A chemical probe of CARM1 alters epigenetic plasticity against breast cancer cell invasion. eLife. 2019;8:e47110.
11. Liu T, et al. Ginsenoside Rg3 regulates DNA damage in non-small cell lung cancer cells by activating VRK1/P53BP1 pathway. 120: Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie; 2019. p. 109483.
12. Liang J, et al. A novel CCCH-zinc finger protein family regulates proinflammatory activation of macrophages. J Biol Chem. 2008;283(10):6337–46.
13. Wang M, et al. Identification of a novel tumor suppressor gene p34 on human chromosome 6q25.1. Cancer research. 2007;67(1):93–9.
14. Wawro M, et al., Intact NYN/PIN-Like Domain is Crucial for the Degradation of Inflammation-Related Transcripts by ZC3H12D. Journal of cellular biochemistry, 2017. 118(3): p. 487–498.
15. Zhu M, et al., miR-128-3p serves as an oncogenic microRNA in osteosarcoma cells by downregulating ZC3H12D. Oncology letters, 2021. 21(2): p. 152.
16. Zhang W, Qiao B, Fan J. Overexpression of miR-4443 promotes the resistance of non-small cell lung cancer cells to epirubicin by targeting INPP4A and regulating the activation of JAK2/STAT3 pathway. Pharmazie. 2018;73(7):386–92.
17. Gong J, et al. lncRNA FEZF1–AS1 contributes to cell proliferation, migration and invasion by sponging miR–4443 in hepatocellular carcinoma. Mol Med Rep. 2018;18(6):5614–20.
18. Gao Y, et al. lncRNA MNX1-AS1 Promotes Glioblastoma Progression Through Inhibition of miR-4443. Oncology research. 2019;27(3):341–7.
19. Meerson A, Yehuda H. Leptin and insulin up-regulate miR-4443 to suppress NCOA1 and TRAF4, and decrease the invasiveness of human colon cancer cells. BMC Cancer. 2016;16(1):882.
20. Wu D, et al. Expression Profiling and Cell Type Classification Analysis in Periodontitis Reveal Dysregulation of Multiple lncRNAs in Plasma Cells. Frontiers in genetics. 2020;11:382.
21. de Lima DS, et al. Long noncoding RNAs are involved in multiple immunological pathways in response to vaccination. Proc Natl Acad Sci USA. 2019;116(34):17121–6.
22. Nagy Á, Munkácsy G, Győrffy B. Pancancer survival analysis of cancer hallmark genes. Scientific reports. 2021;11(1):6047.
23. Zheng Y, et al. lnCAR: A Comprehensive Resource for lncRNAs from Cancer Arrays. Cancer research. 2019;79(8):2076–83.
24. Tabula MC, et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. Nature. 2018;562(7727):367–72.

25. Lambrechts D, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. Nature medicine. 2018;24(8):1277–89.
26. Song Q, et al. Dissecting intratumoral myeloid cell plasticity by single cell RNA-sEq. Cancer medicine. 2019;8(6):3072–85.
27. Zilionis R, et al. Single-Cell Transcriptomics of Human and Mouse Lung Cancers Reveals Conserved Myeloid Populations across Individuals and Species. Immunity. 2019;50(5):1317–34.e10.
28. Newman AM, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nature biotechnology. 2019;37(7):773–82.
29. Kim N, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nature communications. 2020;11(1):2285.
30. Wu TD, et al. Peripheral T cell expansion predicts tumour infiltration and clinical response. Nature. 2020;579(7798):274–8.
31. Lee CM, et al. UCSC Genome Browser enters 20th year. Nucleic acids research. 2020;48(D1):D756–61.
32. Sun D, et al. TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. Nucleic acids research. 2021;49:D1420–30.
33. Vasaikar SV, et al. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. Nucleic acids research. 2018;46:D956–63.
34. Zhou Y, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature communications. 2019;10(1):1523.
35. Li T, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer research. 2017;77(21):e108–10.
36. Shen T, Wang M, Wang X. Identification of Prognosis-related Hub RNA Binding Proteins Function through Regulating Metabolic Processes in Tongue Cancer. J Cancer. 2021;12(8):2230–42.
37. Zhu M, et al., miR-128-3p serves as an oncogenic microRNA in osteosarcoma cells by downregulating ZC3H12D. Oncology letters, 2021. 21(2): p. 152.
38. Vizoso M, et al. Aberrant DNA methylation in non-small cell lung cancer-associated fibroblasts. Carcinogenesis. 2015;36(12):1453–63.
39. Huang S, et al. The putative tumor suppressor Zc3h12d modulates toll-like receptor signaling in macrophages. Cellular signalling. 2012;24(2):569–76.
40. Emming S, et al. A molecular network regulating the proinflammatory phenotype of human memory T lymphocytes. Nature immunology. 2020;21(4):388–99.
41. He Y, et al. Classification of triple-negative breast cancers based on Immunogenomic profiling. Journal of experimental clinical cancer research: CR. 2018;37(1):327.
42. von Wenserski L, et al. SLAMF receptors negatively regulate B cell receptor signaling in chronic lymphocytic leukemia via recruitment of prohibitin-2. Leukemia. 2021;35(4):1073–86.

43. Wawro M, et al., ZC3H12B/MCPIP2, a new active member of the ZC3H12 family. RNA (New York, N.Y.), 2019. 25(7): p. 840–856.
44. Wawro M, et al., Intact NYN/PIN-Like Domain is Crucial for the Degradation of Inflammation-Related Transcripts by ZC3H12D. Journal of cellular biochemistry, 2017. 118(3): p. 487–498.
45. Costantini A, et al., Plasma Biomarkers Screening by Multiplex ELISA Assay in Patients with Advanced Non-Small Cell Lung Cancer Treated with Immune Checkpoint Inhibitors. Cancers, 2020. 13(1).
46. Skundric DS, et al. Emerging role of IL-16 in cytokine-mediated regulation of multiple sclerosis. Cytokine. 2015;75(2):234–48.
47. Gordiienko I, et al., SLAMF1/CD150 in hematologic malignancies: Silent marker or active player? Clinical immunology (Orlando, Fla.), 2019. 204: p. 14–22.
48. Mantsoki A, Devailly G, Joshi A. Gene expression variability in mammalian embryonic stem cells using single cell RNA-seq data. Comput Biol Chem. 2016;63:52–61.
49. Wang L, et al. Myeloid Cell-associated Resistance to PD-1/PD-L1 Blockade in Urothelial Cancer Revealed Through Bulk and Single-cell RNA Sequencing. Clinical cancer research:. an official journal of the American Association for Cancer Research; 2021.
50. Yang L, et al. Development and validation of a prediction model for lung adenocarcinoma based on RNA-binding protein. Annals of translational medicine. 2021;9(6):474.

Figures

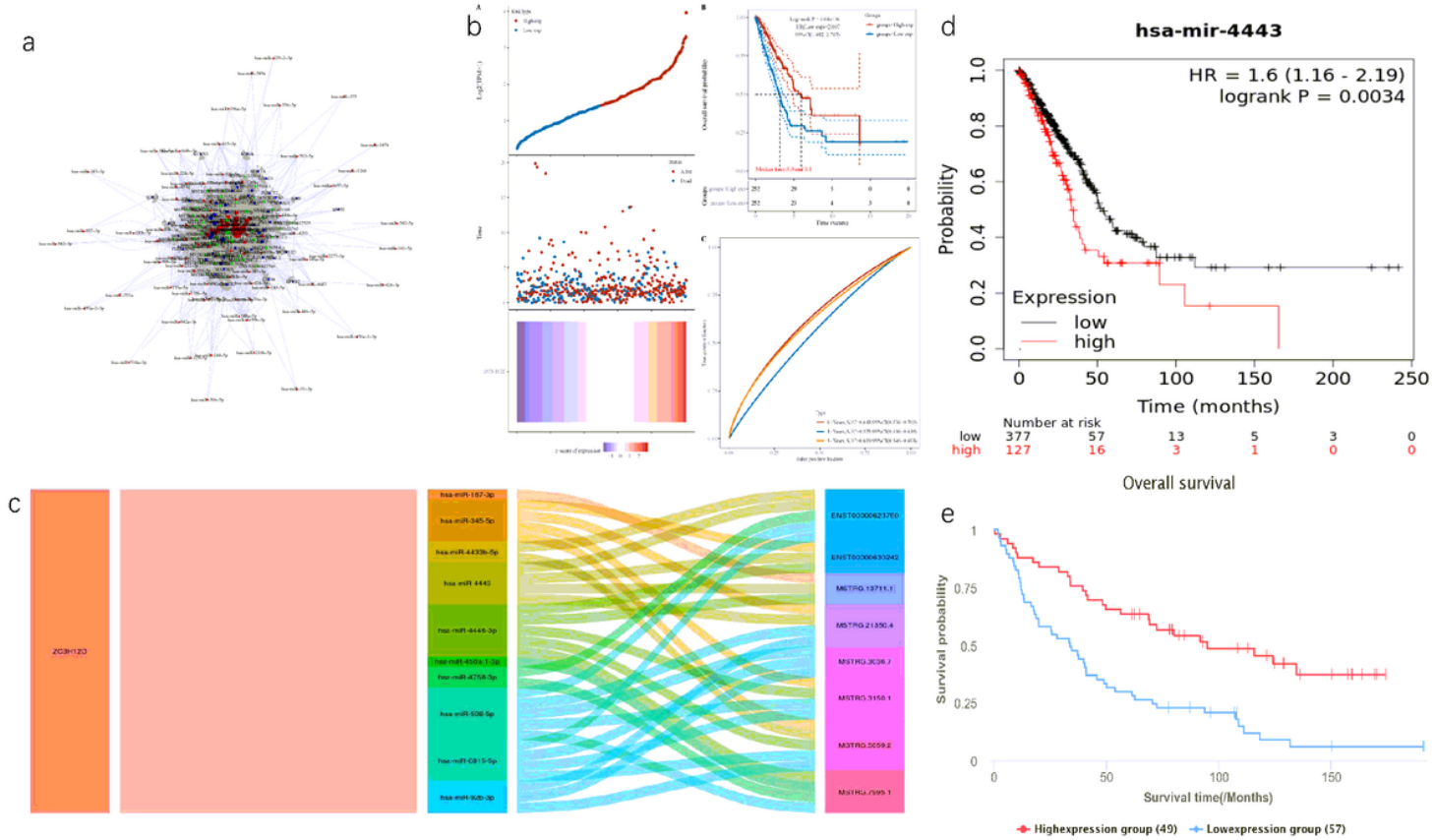

Figure 1

The lncRNA-miRNA-mRNA ceRNA networks constructed (Figure 1a). The green node represents lncRNAs; the red node represents miRNAs; the blue node represents mRNAs. Survival analysis. The ZC3H12D, GNAO1, KSR2, SBK1 and SLIT3 were identified to be significant associated with OS (Figure 1b). The ceRNA network constructed by ZC3H12D and the associated miRNAs and lncRNAs(Figure 1c). High expression of hsa-miR-4443 was not conducive to the prognosis of patients with LUAD (Figure 1d), but high expression of ENST00000630242 was beneficial to the prognosis of patients with LUAD (Figure 1e).

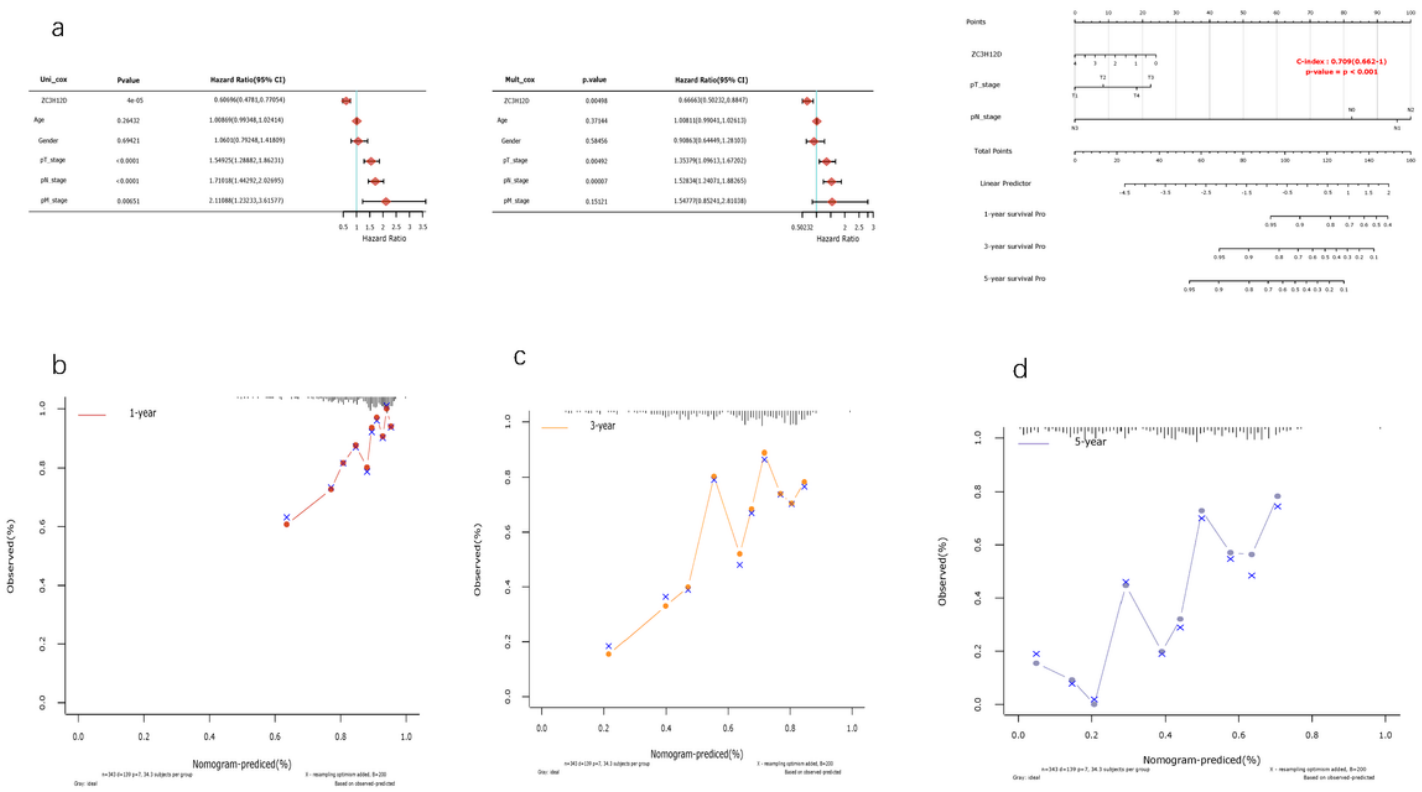

Figure 2

The Cox regression models. ZC3H12D was beneficial to the overall survival of patients in both univariate and multivariate Cox regression models (Figure 2a and Figure 2b). A higher score indicated a higher risk of death (Figure 2c). The calibration plot showed that the nomogram predicting 1-year OS was relatively more accurate than the nomogram predicting 1-year OS,2-year OS and 3-year OS of LUAD patients (Figure 2d).

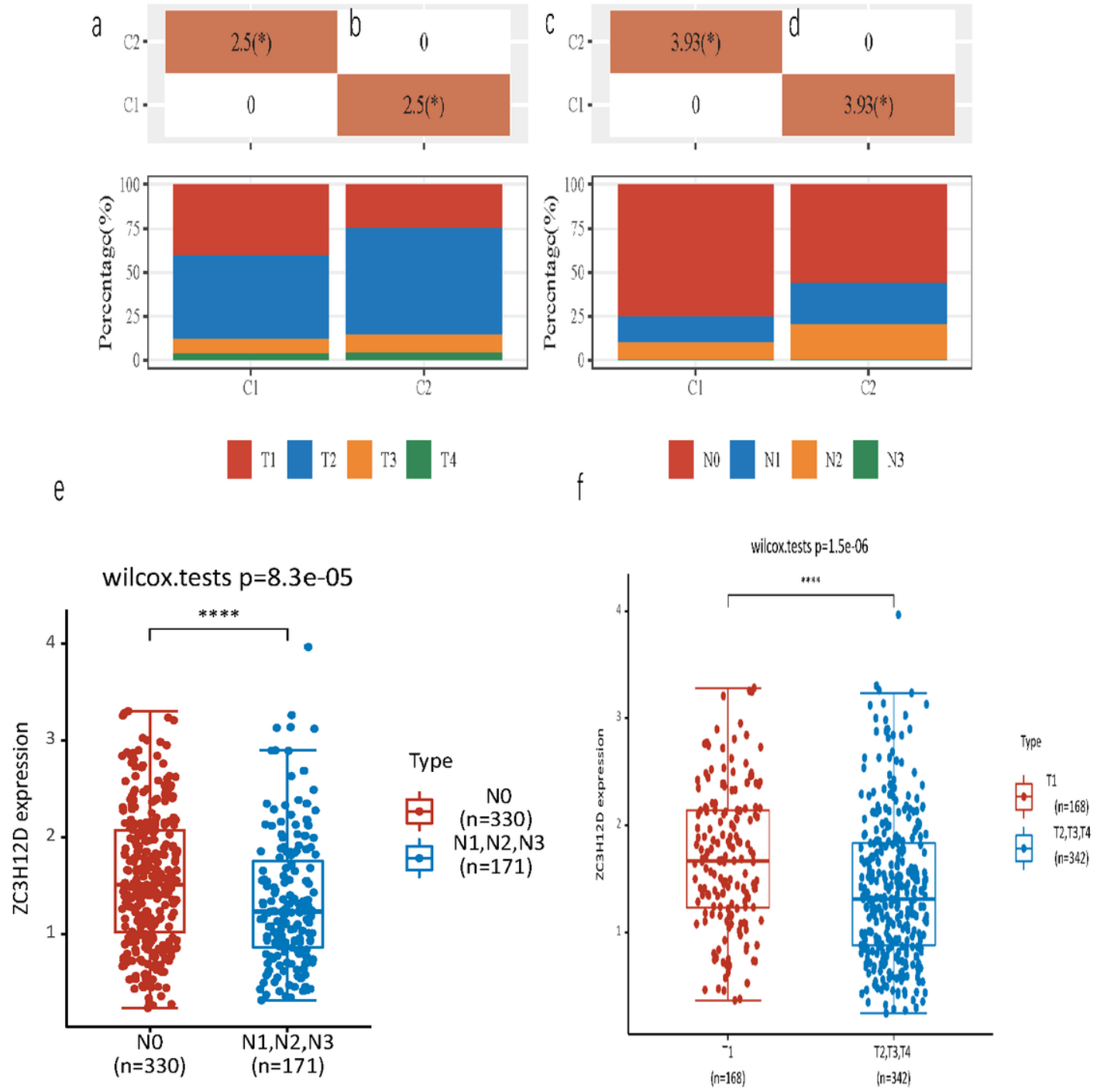

Figure 3

Chi-square test was used to explore the relationship between ZC3H12D expression and clinical factors associated with prognosis. In group C1, the expression of ZC3H12D was significantly higher than the median (Figure 3a), while in group C2 it was significantly lower than the median. and Figure 3b). The correspondent bar chart exhibited that the proportion of ZC3H12D high expression in T1 and N0 was significantly higher than that of ZC3H12D low expression (Figure 3c and Figure 3d). In N0 group, the

ZC3H12D expression was significantly higher than another group (Figure 3e). The ZC3H12D expression was relatively high in T1 phase than the T2,T3and T4 (Figure 3f).

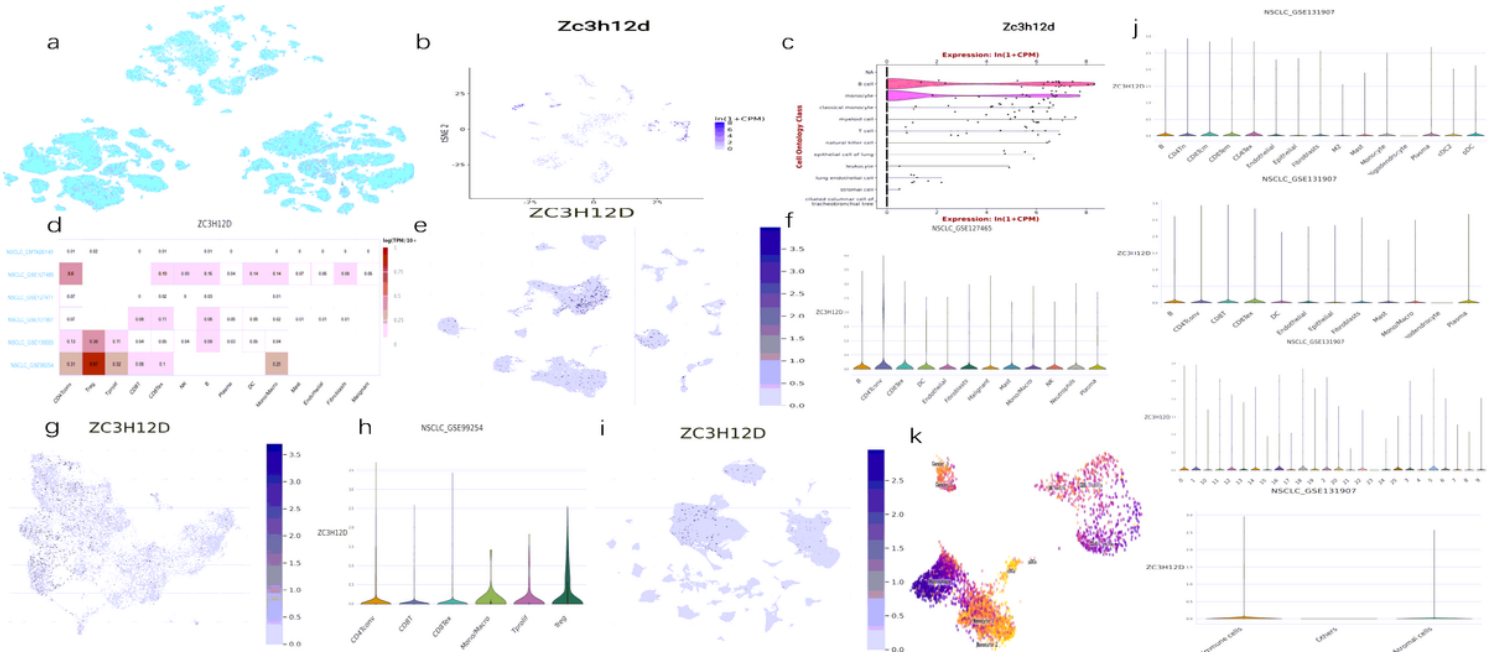

Figure 4

The ZC3H12D was mainly expressed in immune cells,which can be found in both human(Figure 4a and Figure 4c) and mouse(Figure 4b) lung specimens. ZC3H12D expression displayed heterogeneity in different clusters of cells in different NSCLC datasets (Figure 4d). ZC3H12D was more abundantly expressed in CD4 T cells using GSE127465 and in Treg cells using GSE99254 (Figure 4e-h). Another scRNA-seq dataset from GSE131907 indicating that the ZC3H12D was relatively highly expressed in CD8 T cells, B cells and CD4 T cells (Figure 4i and Figure 4j). The ZC3H12D expression in macrophages was significantly increased, and higher than that in CD4 and CD8 T cells (Figure 4k).

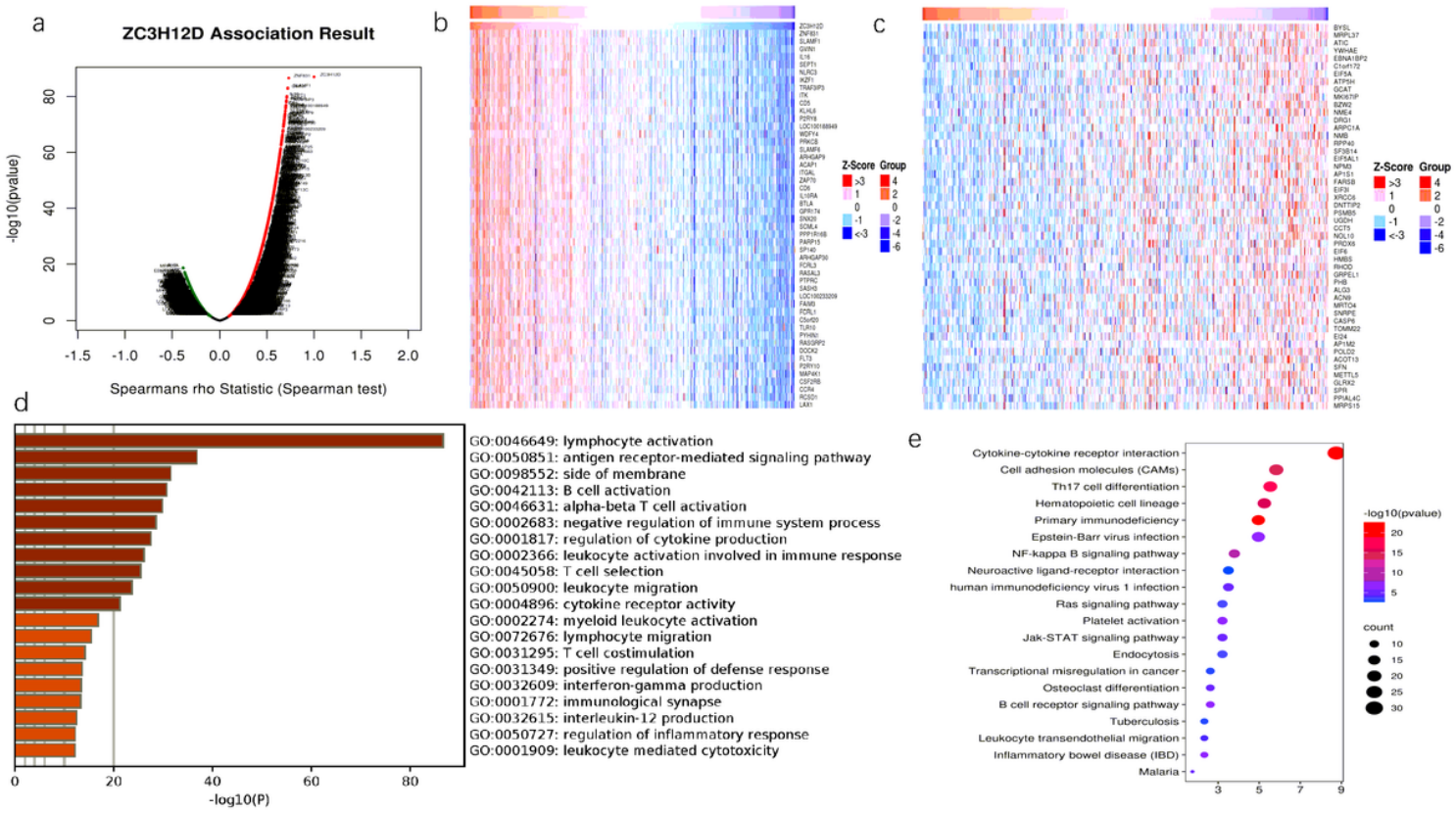

Figure 5

Spearman correlation test. The the volcanic map shows that a total of 12,201 genes were positively correlated with ZC3H12D expression, while 7,786 genes were negatively correlated with ZC3H12D expression(Figure 5a). The top 50 positively correlated significant genes(Figure 5b) and the top 50 negatively correlated significant genes(Figure 5c) were screened and plotted as heat maps. The genes were mainly enriched in the immune function of lymphocyte activation, antigen receptor-mediated signaling pathway, B cell activation and alpha-beta T cell activation (Figure 5d). The co-expressed genes also participated in many pathways, such ascytokine-cytokine receptor interaction, cell adhesion molecules and Th17 cell differentiation (Figure 5e).

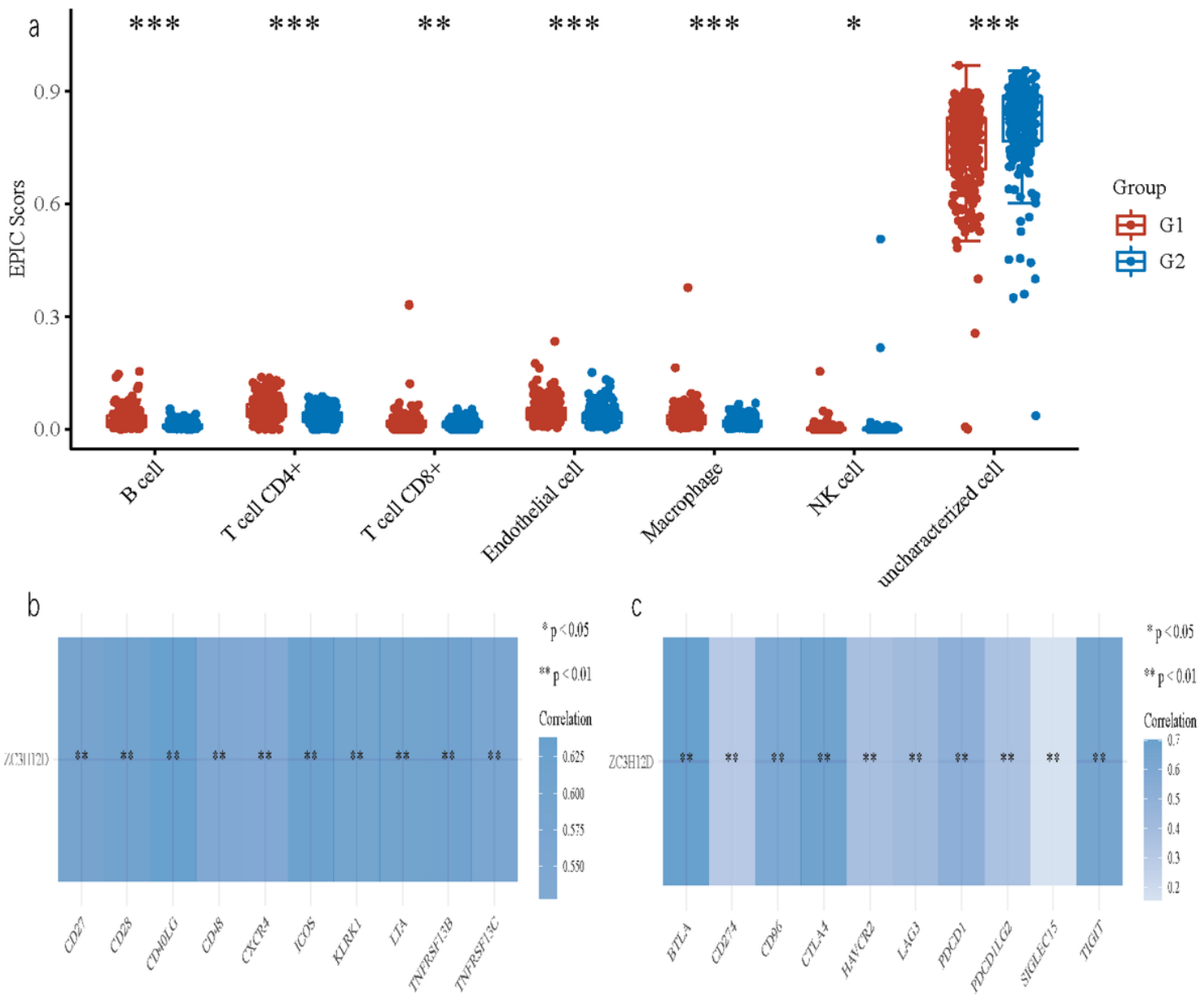

Figure 6

The experimental peritoneal cancer index (EPCI) algorithm showed that six major immune cells showed a trend: the score of the G1 group with high expression was higher than the score of the G2 group with low gene expression, interestingly, in the undefined cells, the score of the G1 group with high gene expression was relatively lower than that of the G2 group with low gene expression (Figure 6a). 10 immunostimulators (CD27,CD28, CD40LG, CD48, CXCR4, ICOS, KLRK1, LTA, TNFRSF13B, TNFRSF13C) with statistical significance (Figure 6b). The expression of ZC3H12D was related to the 10 immunoinhibitors, which were statistically significant, but the correlation between the immune checkpoint SIGLEC15 and ZC3H12D expression is relatively weak (Figure 6c).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- lncRNAandmRNAceRNAnetworklisttop.xlsx
- supplement1.tiff
- NSCLCEMTAB6149pie.png
- NSCLCEMTAB6149umapCelltypecurated.png
- NSCLCGSE99254umapCelltypecurated.png
- NSCLCGSE117570umapCelltypecurated.png
- NSCLCGSE127465umapCelltypecurated.png
- NSCLCGSE131907barplot.png
- NSCLCGSE131907pie.png
- NSCLCGSE131907umapCelltypecurated.png
- NSCLCGSE139555umapCelltypecurated.png
- NSCLCGSE127471umapCelltypecurated.png
- EnrichmentGOColorByPValue.pdf
- EnrichmentGOColorByCluster.pdf