Original Article

Identification and validation of diagnostic alternative splicing events in tumor-educated platelets for non-small cell lung cancer in patients with ground-glass opacity: a multicenter study

Mengqi Shao[1,2#^], Wanwen Li[3#^], Jieming Cao[1,2], Li Wang[1,2], Shenglong Xie[3], Yan Hu[1,2], Gang Feng[3], Feredun Azari[4], Luca Bertolaccini[5], Wenliang Liu[1,2*], Bin He[3*]

[1]Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, Changsha, China; [2]Hunan Key Laboratory of Early Diagnosis and Precision Treatment of Lung Cancer, The Second Xiangya Hospital, Central South University, Changsha, China; [3]Department of Thoracic Surgery, Sichuan Provincial People's Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; [4]Heart, Vascular, and Thoracic Institute at the Cleveland Clinic Foundation, Cleveland, OH, USA; [5]Department of Thoracic Surgery, IEO, European Institute of Oncology IRCCS, Milan, Italy

Contributions: (I) Conception and design: M Shao, W Li, W Liu; (II) Administrative support: B He, L Wang; (III) Provision of study materials or patients: G Feng, S Xie, Y Hu; (IV) Collection and assembly of data: W Li, M Shao; (V) Data analysis and interpretation: M Shao, W Li, J Cao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

[*]These authors contributed equally to this work.

Correspondence to: Bin He, MD. Department of Thoracic Surgery, Sichuan Provincial People's Hospital, School of Medicine, University of Electronic Science and Technology of China, No. 32, Section 2, 1st Ring Road, Qingyang District, Chengdu 610000, China. Email: sy99426@sina.com; Wenliang Liu, MD. Department of Thoracic Surgery, The Second Xiangya Hospital, Central South University, No. 139 Renmin Middle Road, Changsha 410011, China; Hunan Key Laboratory of Early Diagnosis and Precision Treatment of Lung Cancer, The Second Xiangya Hospital, Central South University, Changsha, China. Email: liuwenliang@csu.edu.cn.

Background: The diagnostic potential of tumor-educated platelets (TEPs) across various cancer types has gained increasing recognition; however, the relationship between alternative splicing (AS) events in TEPs and tumor development remains understudied. Early detection of non-small cell lung cancer (NSCLC) in ground-glass opacities (GGOs) is critical for improving patient outcomes, yet current methods lack sufficient accuracy. Our research identified diagnostic-related alternative splicing events (DASEs) in TEPs, revealing several promising biomarkers for NSCLC, specifically in patients presenting with GGOs.

Methods: Patients with GGOs from two hospitals were prospectively enrolled [Hospital 1-Platelet (H1-P) and Hospital 2-Tissue (H2-T) in the validation cohort; Hospital 2-Platelet (H2-P) in the test cohort]. Benign/malignant diagnosis of GGOs was confirmed by pathological examination according to the World Health Organization (WHO) classification. TEPs from the H1-P cohort were collected for transcriptome sequencing and AS analysis. Chi-square tests, least absolute shrinkage and selection operator (LASSO) regression analysis, and protein-protein interaction (PPI) network were used for the preliminary screening of DASEs. Pathological tissue from the H2-T cohort was collected to validate the diagnostic efficacy of hub DASEs in NSCLC against the pathological gold standard. Moreover, TEPs from the H2-P cohort were used to assess the predictive performance of hub DASEs for GGOs using receiver operating characteristic (ROC) curves. Decision curve analysis (DCA) was used to determine whether diagnosing NSCLC in the GGOs population via hub DASEs could benefit patients.

Results: A total of 285 patients with GGOs were enrolled, including 151 NSCLC and 128 inflammatory nodules confirmed by pathological examination. Thirteen DASEs were screened with the chi-square test and LASSO regression analysis to identify diagnostic TEP AS markers for GGOs NSCLC. The PPI network identified four hub diagnostic-related alternative splice genes (DASGs) (*TMEM219*, *MPV17*, *FIBP*, and *VPS28*). Pathological tissues and platelets were collected to validate the four hub DASEs of these four hub DASGs. MXE-32112-*TMEM219* yielded an area under the curve (AUC) of 0.82 [95% confidence interval (CI): 0.729–0.902], with a sensitivity of 83.33% and a specificity of 80.00%; RI-3259-*VPS28* yielded an AUC

^ ORCID: Mengqi Shao, 0000-0002-6554-5381; Wanwen Li, 0009-0003-7272-4673.

of 0.77 (95% CI: 0.677–0.870) with a sensitivity of 93.33% and a specificity of 78.33%; and RI-3641-*MPV17* yielded an AUC of 0.82 (95% CI: 0.728–0.901) with a sensitivity of 90.00% and a specificity of 80.00%. The DCA results suggested that using hub DASEs in diagnosing NSCLC in individuals with GGOs could provide benefits.

Conclusions: The specific diagnostic AS events (MXE-32112-*TMEM219*, RI-3259-*VPS28*, and RI-3641-*MPV17*) identified in TEP samples demonstrated high sensitivity and specificity for diagnosing NSCLC in patients with GGOs. These findings suggest that TEP-related AS events may serve as non-invasive biomarkers to guide biopsy decisions for NSCLC in GGOs, reducing unnecessary procedures.

Keywords: Tumor-educated platelets (TEPs); alternative splicing (AS); ground-glass opacities (GGOs); non-small cell lung cancer (NSCLC); diagnostic marker

Submitted Mar 12, 2025. Accepted for publication Apr 17, 2025. Published online Apr 27, 2025.
doi: 10.21037/tlcr-2025-287
View this article at: https://dx.doi.org/10.21037/tlcr-2025-287

Introduction

With the integration of computed tomography (CT) for screening in high-risk patient populations in recent years, the detection rate of pulmonary ground-glass opacities (GGOs) has gradually increased (1). A GGO is an unclear, dense shadow area with a maximum diameter of <3 cm that still demonstrates bronchial or pulmonary vascular structures on high-resolution CT (2). GGOs include benign and malignant types. The common pathological types of benign GGOs include fibrous focal hyperplasia, alveolar epithelial hyperplasia, and organizing pneumonia (2-4). The common pathological types of malignant GGOs include atypical adenomatous hyperplasia, adenocarcinoma *in situ*, and minimally invasive adenocarcinoma (2-5). CT images can elucidate the diagnosis of benign and malignant GGOs according to the proportion of solid components (6). However, the gold standard for diagnosing GGOs remains based on pathological examination after surgical resection (7). The early diagnosis of cancer is conducive to early treatment and improves patient outcomes. A simpler, more efficient, and less invasive examination method could improve the accuracy of distinguishing between benign and malignant GGOs before surgery, potentially reducing unnecessary surgical resections and contributing to more appropriate diagnosis and treatment plans for patients with GGOs.

Platelets, generated by megakaryocytes in the bone marrow hematopoietic tissue, are the second most abundant enucleated cells in the blood circulation after red blood cells (8). A previous study has found that platelets are primarily involved in hemostasis processes, including thrombus formation, vascular wall repair, and wound healing (9). Recent research findings demonstrate that platelets are essential to cancer progression and growth (10-12). In the bidirectional tumor-platelet interaction, platelets respond systematically and locally to cancer cells and continuously absorb and enrich free proteins,

Highlight box

Key findings

- Three specific diagnostic alternative splicing events (DASEs), namely MXE-32112-*TMEM219*, RI-3259-*VPS28*, and RI-3641-*MPV17*, were identified in tumor-educated platelets (TEPs) for the diagnosis of non-small cell lung cancer (NSCLC) in patients with ground-glass opacities (GGOs). These DASEs showed high diagnostic performance, with an area under the curve values ranging from 0.77 to 0.82, achieving sensitivities between 83.33% and 93.33% and specificities between 78.33% and 80.00%. Decision curve analysis indicated that applying these biomarkers in the GGOs population could offer clinical benefits in detecting early-stage lung cancer.

What is known and what is new?

- TEPs have been regarded as promising diagnostic biomarkers in NSCLC. Alternative splicing (AS) plays a crucial role in tumorigenesis and has the potential to be utilized as a diagnostic tool. Traditional imaging techniques alone cannot reliably differentiate between malignant and benign GGOs.
- This study examined the diagnostic potential of DASEs in TEPs for NSCLC detection among GGOs. The application of these AS markers demonstrated excellent diagnostic performance, providing greater accuracy than conventional methods, especially in the case of GGOs.

What is the implication, and what should change now?

- These hub DASEs may represent valuable tools for the diagnostic toolkit for NSCLC, especially in the early stages where diagnosis and differentiation are often difficult. These biomarkers could lay the foundation for noninvasive diagnostic assays in clinical practice, facilitating earlier intervention and improved patient outcomes.

Table 1 The clinical characteristics of the enrolled patients

Characteristics	Exploration cohort (n=125)		Validation cohort (n=60)		Test cohort (n=100)		P value
	NSCLC (n=59)	Inflammatory nodule (n=66)	NSCLC (n=30)	Inflammatory nodule (n=30)	NSCLC (n=62)	Inflammatory nodule (n=38)	
Gender							<0.001
Male	25 (20.00)	50 (40.00)	13 (21.69)	11 (18.33)	18 (18.00)	15 (15.00)	
Female	34 (27.20)	16 (12.80)	17 (28.33)	19 (31.67)	44 (44.00)	23 (23.00)	
Age (years)	53.85±10.75	60.40±10.22	57.60±10.29	56.27±11.14	58.90±10.17	58.82±10.85	0.44
Smoking status							<0.001
Current	22 (17.60)	40 (32.00)	3 (5.00)	7 (11.67)	10 (10.00)	5 (5.00)	
Former	8 (6.40)	14 (11.20)	17 (28.33)	9 (15.00)	21 (21.00)	14 (14.00)	
Never	29 (23.20)	12 (9.60)	10 (16.67)	14 (23.33)	31 (31.00)	19 (19.00)	

Data are presented as n (%) or mean ± SD. NSCLC, non-small cell lung cancer; SD, standard deviation.

nucleic acids, vesicles, and particles (10-14), which can lead to changes in RNA and proteomic expression profiles (15-17). Therefore, platelets are referred to as tumor-educated platelets (TEPs) due to their altered functions in the presence of tumors (11,18).

TEPs have a variety of roles: they interact with neutrophils, monocytes, and macrophages to stimulate and regulate the function of such immune cells (19,20). TEPs also encapsulate circulating tumor cells, protecting them from immunodetection and the shear stress that could be generated by natural killer cells in the bloodstream (21). In addition, factors released by TEPs enhance the adhesion of tumor cells to endothelial cells and fibrinogen through extracellular vesicles and other pathways, inducing epithelial-mesenchymal transition and facilitating the distant metastasis of cancer (5,22). The TEP-RNA profile has been identified as a diagnostic factor independent of age, smoking, and other predisposing factors in 18 different types of cancer (10,11), including esophageal cancer (23), breast cancer (24), and pancreatic cancer (25). The RNA profile of TEPs in lung cancer has also been significantly altered. Liquid biopsy of the TEP-RNA profile can facilitate the early detection of lung cancer and monitor the dynamic changes of bystander molecules (26), which may improve early non-small cell lung cancer (NSCLC) screening accuracy and potentially prolong patient survival (27-29).

The differential diagnosis of GGOs remains challenging due to overlapping radiological features between benign inflammatory nodules and NSCLC. Unmet needs from previous studies intended to be addressed by this study include the lack of validated noninvasive biomarkers for NSCLC in GGOs, the lack of AS signature studies in TEP, and the need for prospective validation in early disease. We hypothesized that the TEP-AS signature could provide sensitive and specific detection of NSCLC in GGOs. By integrating transcriptomic analysis, machine learning, and pathological validation, this study establishes the first prospective evidence for TEP-AS as a noninvasive diagnostic tool in this population. We present this article in accordance with the TRIPOD reporting checklist (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-2025-287/rc).

Methods

Clinical data and specimen acquisition

The inclusion criteria for the clinical trials included patients age 18 to 80 years, an imaging-based diagnosis of pulmonary GGOs, no history of other malignancies or lung-related diseases, and previous surgical resection of pulmonary nodules. The final diagnosis of all patients was confirmed by postoperative pathological examination as the gold standard (30). This study collected patients' names, gender, age, smoking history, postoperative pathological results, and other relevant information. The detailed information for all included patients is presented in *Table 1*.

All blood samples for the exploration cohort were collected at The Second Xiangya Hospital of Central South University [Hospital 1-Platelet (H1-P)] for platelet extraction and RNA sequencing (RNA-seq). Tissue samples

for the validation cohort [referred to as Hospital 2-Tissue (H2-T)] and blood samples for the test cohort [referred to as Hospital 2-Platelet (H2-P)] were obtained from Sichuan Provincial People's Hospital to verify the differential expression and diagnostic efficacy of hub diagnostic-related alternative splicing events (DASEs).

The tissue specimens were approximately 5 mm × 5 mm × 5 mm in size, and the blood samples were approximately 8 mL in volume. Tissue specimens were stored in 1.5-mL cryovials (Corning Life Sciences, Tewksbury, MA, USA) at −80 ℃ until further analysis. Blood samples were stored in 10-mL ethylene diamine tetra acetic acid (EDTA)-coated anticoagulant tubes (BD Biosciences, Franklin Lakes, NJ, USA) at 4 ℃, and platelets were extracted within 48 hours for subsequent experiments.

This study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. This study was approved by the medical ethics committees of both The Second Xiangya Hospital of Central South University (No. 2020-107) and Sichuan Provincial People's Hospital (No. 2025-102). All data and specimens were collected with the patient's informed consent, and relevant consent forms were signed before participation.

Platelet extraction

The blood specimens were slowly centrifuged at 200 ×g for 20 minutes, with the centrifuge gradient set at the lowest setting to protect cell viability. At the end of centrifugation, the specimen was divided into three layers: the upper plasma layer containing plasma proteins and so on, the middle layer which is relatively thin and contains white blood cells and platelets, and the lower layer mainly made up of red blood cells. The transparent liquid on the top layer, which contains a significant amount of platelets along with other plasma components, was transferred to a new 15-mL centrifuge tube and slowly centrifuged at 360 ×g for 10 minutes. This second centrifugation caused the platelets to sediment. After the supernatant was discarded, 1 mL of red blood cell lysate (Haoyang Biological Products Technology Co., Ltd., Tianjin, China) was added to each 15-mL centrifuge tube. After remixing, the mixture was slowly centrifuged at a 400 ×g centrifugal force for 10 minutes. The supernatant was discarded, and 1 mL of TRIzol (Xinjing Biological Reagent Development Co., Ltd., Hangzhou, China) was added to the centrifuge tube, blown and mixed, and immediately stored at −80 ℃ until analysis.

RNA extraction

RNA extraction from all tissue specimens in this study was performed using the Steady Pure Universal RNA Extraction Kit (Accurate Biotechnology Co., Ltd., Changsha, China). Subsequently, 200 µL of Buffer EX (Xinjing Biological Reagent Development Co., Ltd.) was added to the platelet extract and TRIzol mixture and centrifuged at 12,000 ×g for 15 minutes. An additional RNase-free 1.5-mL centrifuge tube was used to add 500 µL of isopropanol, and the transparent upper phase formed by the centrifugation, as mentioned above, was transferred to this centrifuge tube and centrifuged at 12,000 ×g for 15 minutes. The supernatant was then discarded. Then, 1 mL of 75% ethanol (Kelon Chemicals Co., Ltd., Chengdu, China) was added and centrifuged at 7,500 ×g for 5 minutes. The remaining ethanol was removed, and the RNA was allowed to stand at room temperature for 5 minutes to dry. The extracted RNA concentration was measured and recorded using a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA, USA) and stored at −80 ℃.

TEP RNA-seq and AS analysis

Platelet samples from the exploration cohort were subjected to RNA-seq. The detailed RNA-seq procedure has been described elsewhere (29). Cutadapt (31) was used to control the quality of the sequencing results. The following three kinds of sequences were filtered: (I) the sequence with the adapter, (II) sequences with N bases (N base number >5), and (III) more than 50% of the sequences with an essential quality value lower than 20. After filtering, HISAT2 (32) (v. 2.2.1) was used for sequence alignment. Finally, rMATS (33) (v. 4.0) was used to perform AS classification, and differential AS analysis from RNA-seq data AS analysis was performed for reads containing splice sites aligned with skipping exons. In this kind of analysis, AS was divided into five types: skipped exon (SE), mutually exclusive exon (MXE), alternative 5' splice site (A5SS), alternative 3' splice site (A3SS), and retained intron (RI).

Polymerase chain reaction (PCR) experiments and primer design

After RNA extraction from tissues and platelets, the same reaction system and reaction conditions were used for PCR experiments. The principles of primer design were as follows: for RI events, primers were designed for

the included intron fragment, and MXE events, primers were designed for both SEs. Reference sequences for all primers were downloaded from the National Center for Biotechnology Information (NCBI), synthesized by Qingke Biological Co., Ltd. (Beijing, China), and stored at −20 ℃. All primer sequences are listed in the Table S1. PCR experiments were performed on a QuantStudio1 Plus real-time PCR instrument (Thermo Fisher Scientific). The primer concentration was adjusted to 10 μM, and the PCR reaction system was 20 μL. GAPDH was chosen as an internal control for all PCR experiments. The PCR reaction was cycled at 95 ℃ for 30 seconds for predenaturation, followed by 40 cycles at 95 ℃ for 5 seconds and 60 ℃ for 30 seconds for denaturation and annealing. Finally, the dissociation step was carried out. The dissolution curve was observed at the end of the reaction, and the data were analyzed with the $2^{-\Delta\Delta CT}$ method. All experiments were repeated three times.

Statistical analysis

The chi-square test was used to assess the association between AS events and the diagnosis of NSCLC. The degree of freedom was set to 1, and the chi-square value of each AS event was calculated. The results from the chi-square test were analyzed via least absolute shrinkage and selection operator (LASSO) regression analysis for machine learning. A regression model that included all identified AS events and an L1 norm penalty term was constructed. The LASSO penalty parameter (λ) was set as the system's default value to control the shrinkage of regression coefficients. The optimal λ value was selected by cross-validation to minimize the model's prediction error. Finally, AS events with nonzero coefficients in the final model were considered significant features and were subsequently identified as AS events associated with NSCLC diagnosis in patients with GGOs. To examine the functional relationships between diagnostic-related alternative splice genes (DASGs), a protein-protein interaction (PPI) network was constructed with the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (https://cn.string-db.org/) and GeneMANIA (https://genemania.org/). The Markov clustering (MCL) algorithm was used to calculate the PPI score of DASGs. The results of PCR experiments were analyzed using an independent *t*-test. Receiver operating characteristic (ROC) curve analysis was performed using the pROC package, with area under the curve (AUC) and 95% confidence interval (CI) calculated via the DeLong method. Sensitivity and specificity were derived at the optimal cutoff determined by Youden's index. All P values were two-sided with $\alpha = 0.05$. Statistical significance was set at $P<0.05$, with a 95% CI. The statistical analyses and data visualizations were performed using R version 4.2.3 (The R Foundation of Statistical Computing) and GraphPad Prism version 9.5.0 (GraphPad Software, La Jolla, CA, USA). Image stitching and annotation were conducted using Adobe Illustrator 2023 (Adobe, San Jose, CA, USA).

Results

AS events occurred in both NSCLC and inflammatory nodules

One hundred twenty-five platelet samples from patients with GGOs (59 NSCLC and 66 inflammatory nodules) were included in the exploration cohort (H1-P) for RNA-seq. The validation cohort (H2-T) included 60 histological samples of the GGOs population (30 NSCLC and 30 inflammatory nodules) to validate the diagnostic markers selected by the exploration cohort. The test cohort (H2-P) included 100 TEP samples from patients with GGOs (62 NSCLC and 38 inflammatory nodules) to test the validation cohort results further. The clinical information of the enrolled patients is shown in *Table 1*.

Via RNA-seq, 4,547 AS events in 2,664 altered genes were screened from the GGOs population, named as differential expressed alternative splice events (DEASEs). Among them, 246 A3SS events occurred in 208 altered genes, 171 A5SS events occurred in 158 altered genes, 2,164 MXE events occurred in 1,011 altered genes, 685 RI events occurred in 507 altered genes, and 1,281 SE events occurred in 780 altered genes. Patients with NSCLC had 2,296 AS events in 1,481 altered genes, including 106 A3SS events in 98 altered genes, 69 A5SS events in 64 altered genes, 898 MXE events in 495 altered genes, 618 RI events in 461 altered genes, and 605 SE events in 363 altered genes. Patients with inflammatory nodules had 2,251 AS events in 1,456 altered genes, including 140 A3SS events in 117 altered genes, 102 A5SS events in 96 altered genes, 1,266 MXE events in 692 altered genes, 67 RI events in 62 altered genes, and 676 SE events in 489 altered genes (*Figure 1*).

Screening of DASEs

Further analysis was conducted to screen out the DASEs. DEASEs between NSCLC and inflammatory nodules were screened from the RNA-seq results of H1-P. Chi-square

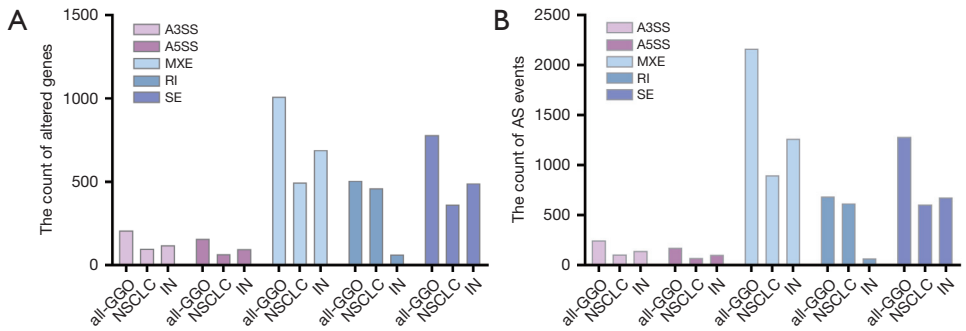

Figure 1 Number of (A) altered genes and (B) alternative splice events in GGOs and NSCLC nodules and IN nodules in the H1-T cohort. H1-T: Hospital 1-Tissue. A3SS, alternative 3' splice site; A5SS, alternative 5' splice site; AS, alternative splicing; GGOs, ground-glass opacities; IN, inflammatory nodule; MXE, mutually exclusive exons; NSCLC, non-small cell lung cancer; RI, retained intron; SE, skipped exon.

tests were performed on DEASEs to assess the association between the AS events of TEPs and the diagnosis of NSCLC. The DEASEs with statistical significance (P<0.05) in the chi-square test were labeled DASE-1s. There were 1,564 DASE-1s (table available at https://cdn.amegroups.cn/static/public/10.21037tlcr-2025-287-1.xlsx), including 101 A3SS, 82 A5SS, 563 MXE, 484 RI, and 334 SE events. The top 100 DASEs in terms of highest incidence were plotted in heat maps (*Figure 2A*), and DASEs appeared to differ between inflammatory and NSCLC nodules.

The H1-P cohort was divided into training and validation populations at a ratio of 7:3 for LASSO regression (*Figure 2B,2C*). LASSO regression identified the 13 most representative DASEs from among the DASE-1s, which were labeled DASE-2s. DASE-2s were visualized and analyzed, with bar plots being drawn (*Figure 2D*). It was found that the 13 DASE-2s had varying degrees of influence on the diagnosis of NSCLC in the GGOs population, with the absolute coefficient value of MXE-11950-*RPL26* being the largest, suggesting that it had the most significant influence. Based on the DASE-2s, a preliminary diagnostic model for NSCLC in the GGOs population was established, and the ROC was drawn. The AUC was calculated to be >0.9, indicating that the diagnostic performance of the diagnostic model was excellent (*Figure 2E*).

The 12 genes corresponding to DASE-2 were considered to be DASGs. A PPI network was visualized to screen out the more closely connected DASGs (*Figure 2F*). Ultimately, four DASGs (*TMEM219*, *FIBP*, *VPS28*, and *MPV17*) with a higher interaction score were selected and considered the hub DASGs. The AS events corresponding to these four hub DASGs in the LASSO analysis were considered to be hub DASEs and included MXE-32112-*TMEM219*, RI-2072-*FIBP*, RI-3259-*VPS28*, and RI-3641-*MPV17*.

Validation of hub DASEs

To evaluate the differential expression of hub DASEs between NSCLC and inflammatory nodules in the GGOs population, 60 tissue samples (cohort H2-T: 30 from patients with NSCLC and 30 from patients with inflammatory nodules) were collected for quantitative real-time PCR. The bar chart and expression heat map (*Figure 3A-3E*) suggested that the abundance of MXE-32112-*TMEM219* (P<0.001), RI-2072-*FIBP* (P<0.001), RI-3259-*VPS28* (P<0.001), and RI-3641-*MPV17* (P<0.0001) in GGOs patients with NSCLC was significantly higher than that in inflammatory nodules. The ROC curve showed that MXE-32112-*TMEM219* (AUC =0.89; 95% CI: 0.792–0.967), RI-2072-*FIBP* (AUC =0.84; 95% CI: 0.725–0.929), RI-3259-*VPS28* (AUC =0.90; 95% CI: 0.818–0.973), and RI-3641-*MPV17* (AUC =0.92; 95% CI: 0.847–0.982) (*Figure 3F*) had high diagnostic efficacy in GGOs tissue samples. In addition, the decision curve analysis (DCA) of the diagnostic performance of these four markers in the validation cohort showed that the decision curve was distant from the two extreme curves (*Figure 3G-3J*). This indicated that applying these four hub DASEs in diagnosing NSCLC in patients with GGOs has potential clinical value and may help reduce unnecessary surgical resection, thereby benefiting patients.

This study aimed to assess a simple and highly available liquid biopsy method for the early diagnosis of NSCLC in patients with GGOs. Therefore, based on the verification results of histological samples, the potential of hub DASEs to diagnose NSCLC in the GGOs population was

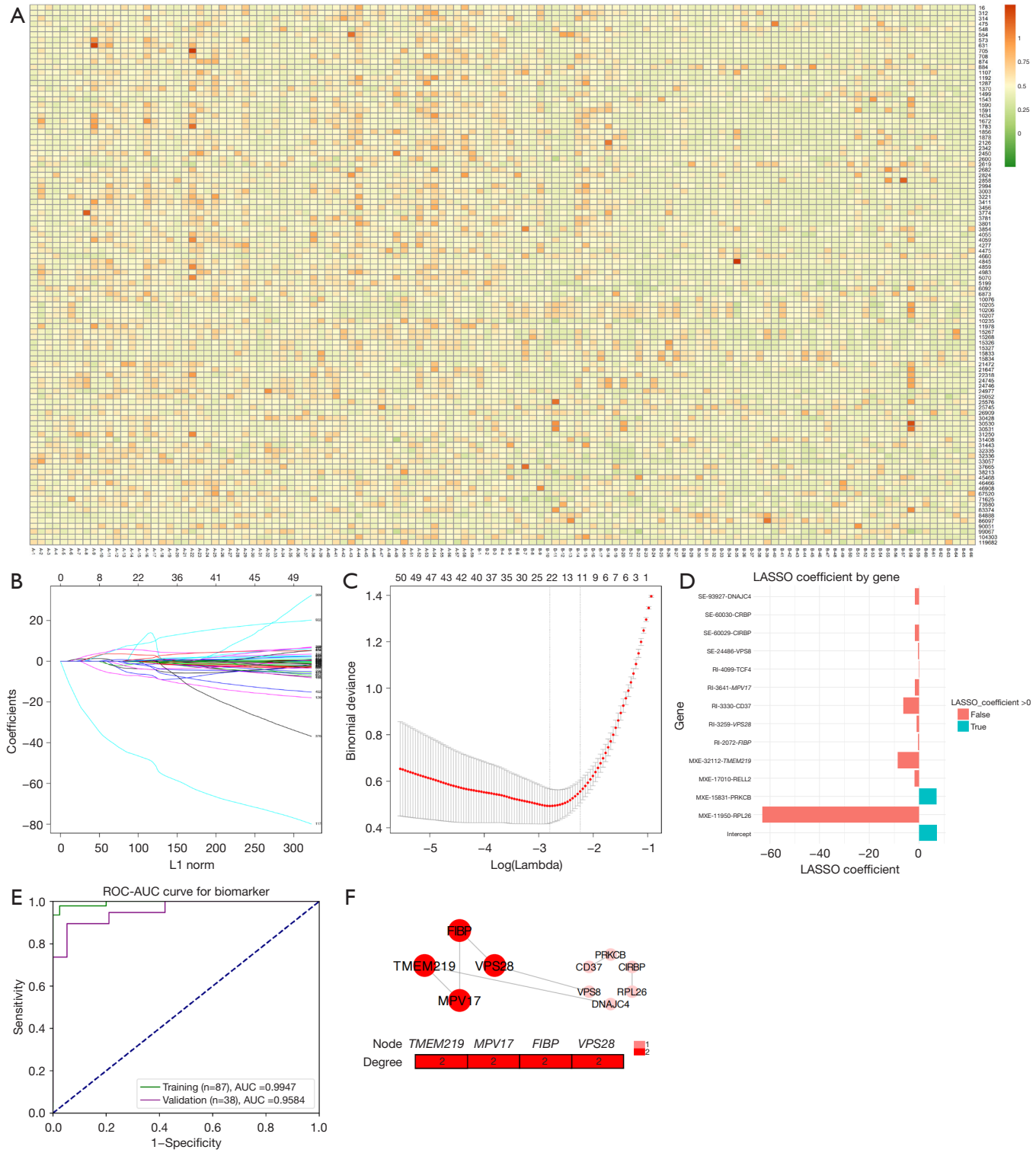

Figure 2 The alternative splicing events related to the diagnosis of NSCLC were screened in the GGOs population. (A) The heatmap of DASE-1s (P<0.05). Selection of NSCLC DASE-1s in GGOs (H1-T cohort) via LASSO regression analysis. (B) Coefficient paths, (C) cross-validation curves, (D) bar plot of DASE-2s, and (E) their ROC AUC (AUC of the training cohort: 0.9947; AUC of the validation cohort: 0.9584). (F) The PPI network of DASGs via Cytoscape. H1-T: Hospital 1-Tissue. AUC, area under the curve; DASE, diagnosis-related

alternative splicing event; DASG, diagnostic-related alternative splice gene; GGOs, ground-glass opacities; LASSO, least absolute shrinkage and selection operator; NSCLC, non-small cell lung cancer; PPI, protein-protein interaction; ROC, receiver operating characteristic.

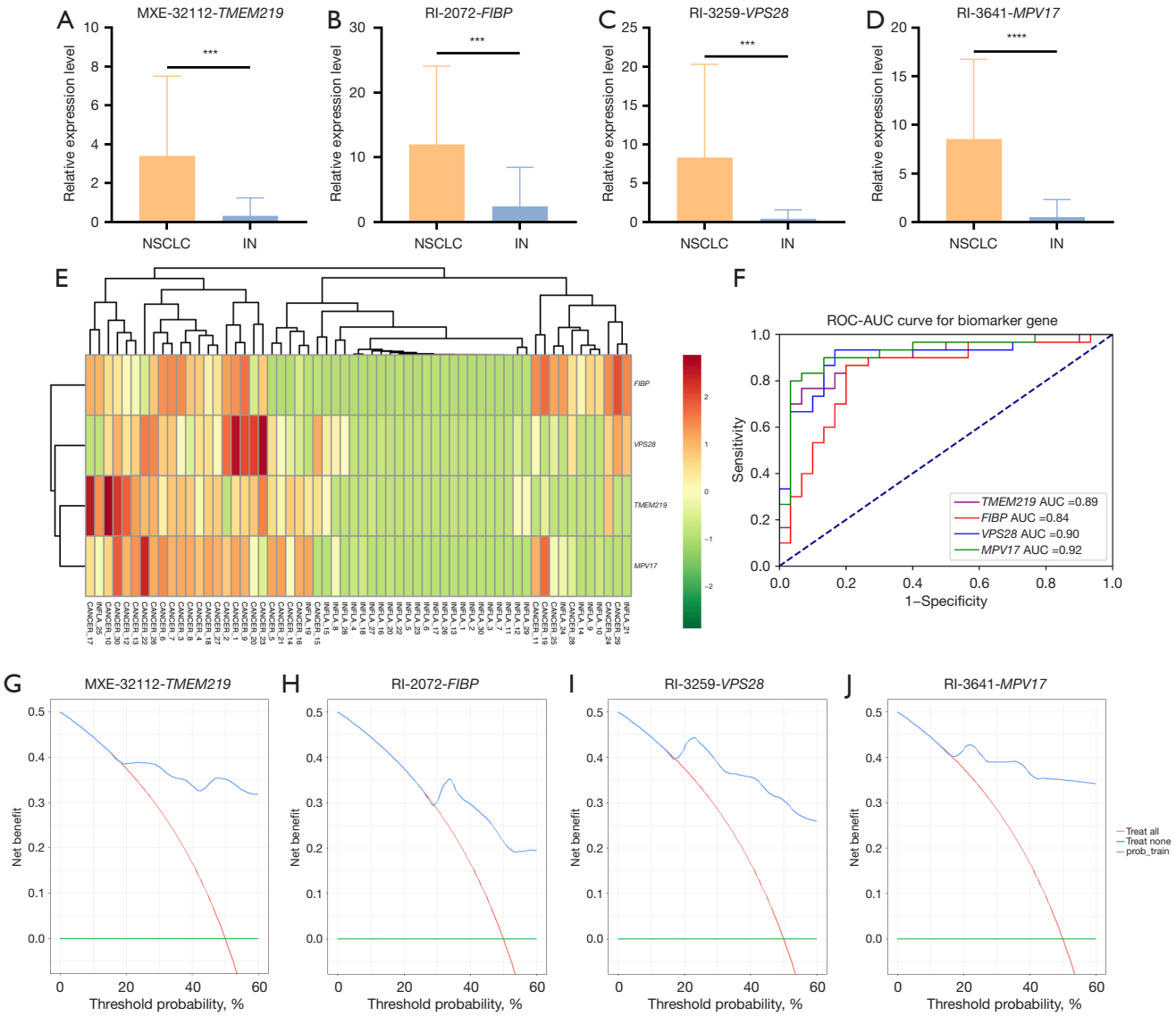

Figure 3 Validation results for tissue samples from the H2-T cohort. Relative expression levels of (A) MXE-32112-*TMEM219*, (B) RI-2072-*FIBP*, (C) RI-3259-*VPS28*, (D) and RI-3641-*MPV17* in NSCLC and inflammatory nodules from the H2-T cohort. (E) Heatmap of MXE-32112-*TMEM219*, RI-2072-*FIBP*, RI-3259-*VPS28*, and RI-3641-*MPV17* in NSCLC and inflammatory nodules from the H2-T cohort. (F) ROC-AUC curves of MXE-32112-*TMEM219* (AUC: 0.89), RI-2072-*FIBP* (AUC: 0.84), RI-3259-*VPS28* (AUC: 0.90), and RI-3641-*MPV17* (AUC: 0.92) in the H2-T cohort. Decision curves of (G) MXE-32112-*TMEM219*, (H) RI-2072-*FIBP*, (I) RI-3259-*VPS28*, (J) and RI-3641-*MPV17* verified the diagnostic performance in the H2-T cohort. H2-T: Hospital 2-Tissue. ***, P<0.001; ****, P<0.0001. AUC, area under curve; IN, inflammatory nodule; MXE, mutually exclusive exons; NSCLC, non-small cell lung cancer; RI, retained intron; ROC, receiver operating characteristic.

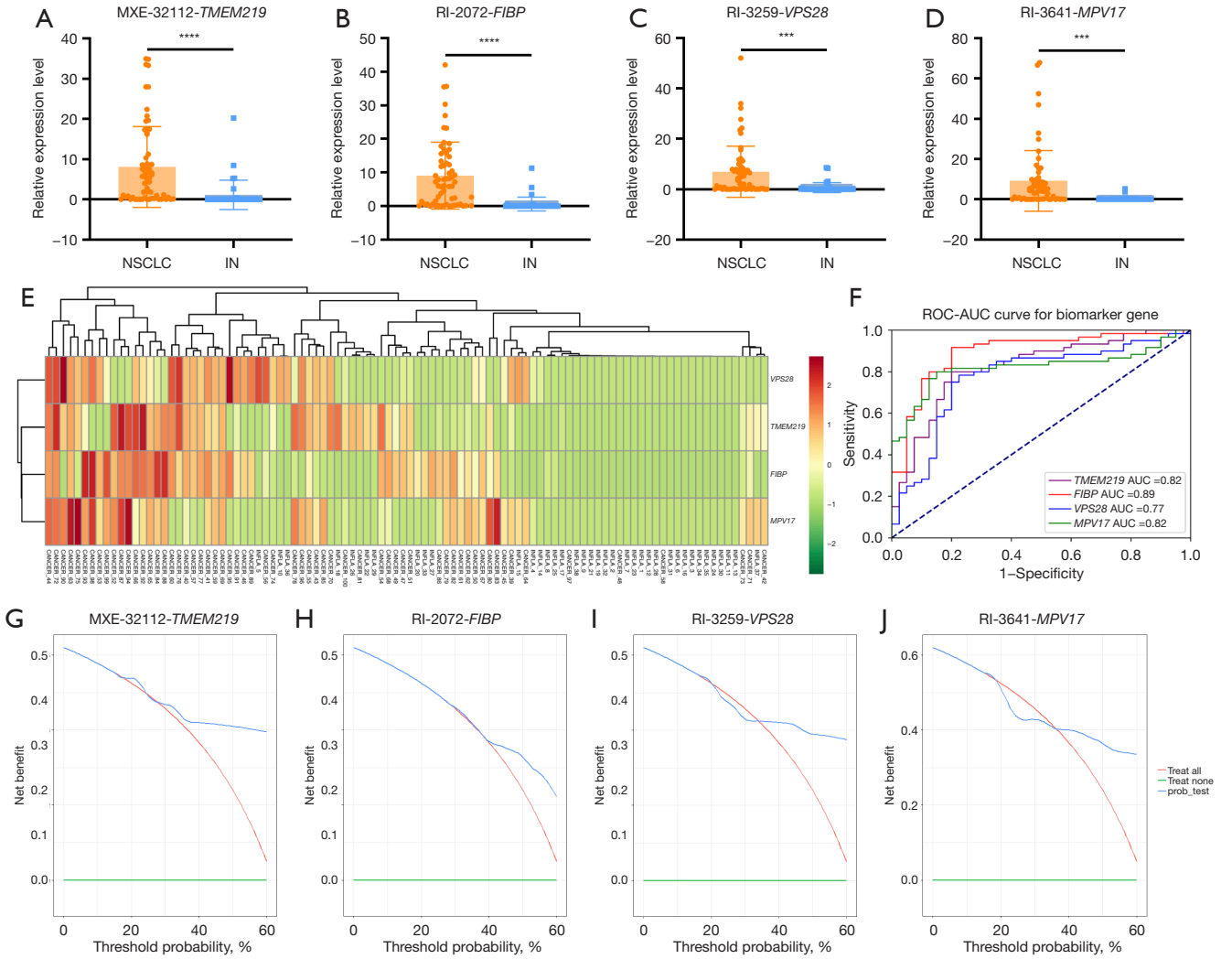

Figure 4 Verification of hub DASEs in TEP samples from the H2-P cohort. (A-D) Bar plots and (E) heatmap of the relative expression levels of (A) MXE-32112-*TMEM219*, (B) RI-2072-*FIBP*, (C) RI-3259-*VPS28*, (D) and RI-3641-*MPV17* in NSCLC and inflammatory nodules from the H2-P cohort. (F) ROC AUC of MXE-32112-*TMEM219* (AUC: 0.82), RI-2072-*FIBP* (AUC: 0.89), RI-3259-*VPS28* (AUC: 0.77), and RI-3641-*MPV17* (AUC: 0.82) in the H2-T cohort. Decision curves of (G) MXE-32112-*TMEM219*, (H) RI-2072-*FIBP*, (I) RI-3259-*VPS28*, and (J) RI-3641-*MPV17* in the H2-T cohort. H2-P: Hospital 2-Platelet; H2-T: Hospital 2-Tissue. ***, P<0.001; ****, P<0.0001. AUC, area under the curve; DASE, diagnosis-related alternative splicing event; IN, inflammatory nodule; MXE, mutually exclusive exons; NSCLC, non-small cell lung cancer; RI, retained intron; ROC, receiver operating characteristic; TEP, tumor-educated platelet.

further verified in TEP samples. Thus, 100 TEP samples (H2-P) from patients with GGOs were prospectively collected before surgery, and the relative expression of hub DASEs in H2-P was tested to predict the postoperative pathological results (38 cases of inflammatory nodules, 62 cases of NSCLC). Sensitivity and specificity were calculated to verify the efficacy of the four hub DASEs in diagnosing patients with NSCLC in the GGOs population. Analysis of the expression levels of MXE-32112-*TMEM219* (P<0.0001) (*Figure 4A*), RI-2072-*FIBP* (P<0.0001) (*Figure 4B*), RI-3259-*VPS28* (P<0.001) (*Figure 4C*), and RI-3641-*MPV17* (P<0.001) (*Figure 4D*) in TEPs in NSCLC and inflammatory nodules, and the bar graph and heat map (*Figure 4E*) were drawn, the results showed that expression levels of these four hub DASEs in NCSCL was significantly higher than that in inflammatory nodules. MXE-32112-

TMEM219 yielded an AUC of 0.82 (95% CI: 0.729–0.902), with a sensitivity of 83.33% and a specificity of 80.00%. RI-2072-*FIBP* yielded an AUC of 0.89 (95% CI: 0.818–0.950), with a sensitivity of 16.67% and a specificity of 91.67%. RI-3259-*VPS28* yielded an AUC of 0.77 (95% CI: 0.677–0.087), with a sensitivity of 93.33% and a specificity of 78.33%. RI-3641-*MPV17* yielded an AUC of 0.82 (95% CI: 0.728–0.901), with a sensitivity of 90.00% and a specificity of 80.00% (*Figure 4F*). Similarly, we performed DCA on the diagnostic efficacy of these four markers in the test cohort and found that the decision curve was distant from the two extreme curves (*Figure 4G-4J*). The greater the area under the decision curve is, the stronger the clinical decision-making power. Therefore, these four hub DASEs may aid in diagnosing NSCLC in patients with GGOs. Thus, given the low sensitivity of RI-2072-*FIBP*, only MXE-32112-*TMEM219*, RI-3259-*VPS28*, and RI-3641-*MPV17* were ultimately considered as diagnostic AS markers for NSCLC in the GGOs population.

Discussion

This study used RNA-seq of TEPs in GGOs to identify NSCLC DASEs. We collected tissues and TEP samples from different hospitals to validate the diagnostic efficacy of DASEs. Three DASEs (MXE-32112-*TMEM219*, RI-3259-*VPS28*, and RI-3641-*MPV17*) identified in this study should be further investigated as potential diagnostic markers and may aid clinicians in determining whether patients with GGOs should undergo surgical resection or further observation.

Tumor cells are known to interact with platelets, potentially altering their RNA profile (34). These TEPs play a significant role in the malignant behavior of tumors, facilitating tumor progression (26,34,35). However, the specific mechanism of how tumor cells "educate" TEPs has not been fully understood. A previous study has suggested that the highly dynamic RNA profile of TEP may be related to the direct uptake (splicing) of circulating messenger RNA (mRNA) by immature platelets in response to external signals (36). Changes in the RNA profile may be due to the production of different mRNA variants from pre-mRNA through AS (37,38). AS can enable a single gene to generate multiple mRNA variants that encode functionally similar or other proteins, facilitating gene diversity and functional adaptation across biological processes. Thus, we hoped to predict the pathological nature of GGOs by characterizing the specific A events produced by TEP.

This study collected 125 TEP samples (59 NSCLC and 66 inflammatory nodules) from the H1-P cohort for RNA-seq and AS analysis. The results suggest differences in AS events in TEPs between NSCLC and inflammatory nodules. More specifically, TEPs in NSCLC nodules tended to have RI AS events more frequently than those in inflammatory nodules. RI events can lead to premature stop codons, allowing the mRNA to be recognized as an abnormal transcript, triggering nonsense-mediated mRNA degradation (NMD) to be identified and degraded by cellular quality-control mechanisms, resulting in decreased functional protein levels (39,40).

Subsequently, three hub DASEs (MXE-32112-*TMEM219*, RI-3259-*VPS28*, and RI-3641-*MPV17*) showed high specificity and sensitivity in the diagnosis of NSCLC in patients with GGOs. Each hub-altered gene plays a significant role in tumor biology. The *TMEM219* gene is a specific cell death receptor of *IGFBP3*, binding to *IGFBP-3R* to release procaspase-8, activating caspase-8-dependent apoptosis (41). The MXE event may affect the binding efficiency of *TMEM219* to *IGFBP-3*, thereby reducing the release of procaspase-8 and inhibiting the initiation of apoptotic signals, leading to weakened cell death signals and promoting tumor survival and proliferation (42). After MXE occurs, the nonfunctional or partially functional TMEM219 protein may be formed, which may lead to partial or complete loss of function of the apoptosis signaling pathway, thus affecting the normal regulation of apoptosis (43). Certain spliceosomes may alter the localization of *TMEM219* on the cell membrane or its interaction with other signaling molecules, activating alternative or atypical signaling pathways and promoting the malignant progression of tumors (43-45). Vacuolar protein sorting-associated protein 28 (*VPS28*) is one of the cytoplasmic proteins of the endosomal sorting complex required for transport I (ESCRT-I) (45-47). The structural or functional alterations caused by the RI event in *VSP28* may interfere with the function of the endosomal system, thereby affecting the normal sorting process of proteins (48). Abnormalities in protein transport function may result in the inability of some specific proteins to be degraded or mislocalized, resulting in accumulation within the cell (46,49,50). *MPV17* is a mitochondrial inner membrane protein responsible for maintaining the stability of mitochondrial DNA (mtDNA) (51). Previous studies have suggested that the deletion of *MPV17* can lead to mtDNA depletion in various cells, increase the abundance of reactive oxygen species (ROS), and promote cell apoptosis, suggesting that

MPV17 plays a protective role in cells (51,52). RI events in *MPV17* may shorten or lead to incomplete translation of the MPV17 protein, disrupting mtDNA replication and stability and compromising mitochondrial function.

We employed tissue and TEP samples for double validation to improve the accuracy of the diagnostic efficacy of TEP-AS diagnostic markers. Additionally, we conducted exploration and validation in multiple centers, reducing potential geographical biases and improving sample diversity and reliability. However, some limitations should still be considered. Although our study contained a relatively large dataset, future studies with larger, more diverse samples are required for further validation. Moreover, previous studies have suggested that the TEP-RNA profile is dynamically changing (53,54), but our study lacked longitudinal data for dynamic detection.

A strength of this study was its multicenter design. TEP samples from the principal study center (H1-P cohort) were collected for transcriptome sequencing and AS analysis. We further screened and confirmed hub DASEs with promising diagnostic potential. Histological and TEP samples were collected at the associated centers (H2-T cohort and H2-P cohort) for double validation.

Conclusions

The final results indicated a high sensitivity and specificity of MXE-32112-*TMEM219*, RI-3259-*VPS28*, and RI-3641-*MPV17* for diagnosing NSCLC in the GGOs population. It is feasible to detect the specific AS markers via liquid biopsy of TEP samples for the early diagnosis of NSCLC in the GGOs population.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-2025-287/rc

Data Sharing Statement: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-2025-287/dss

Peer Review File: Available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-2025-287/prf

Funding: The work was funded by the National Natural Science Foundation of China (grant No. 82172879), the Natural Science Foundation of Hunan Province for Distinguished Young Scholars (grant No. 2022JJ10096), the Key Research and Development Program of Science and Technology Department of Sichuan Province (grant No. 2023YFS0305), and the Natural Science Foundation of Sichuan Province (grant No. 2022NSFSC0780).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tlcr.amegroups.com/article/view/10.21037/tlcr-2025-287/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was conducted in accordance with the Declaration of Helsinki and its subsequent amendments. This study was approved by the medical ethics committees of both The Second Xiangya Hospital of Central South University (No. 2020-107) and Sichuan Provincial People's Hospital (No. 2025-102). All data and specimens were collected with the patient's informed consent, and relevant consent forms were signed before participation.

References

1. Zhang Y, Fu F, Chen H. Management of Ground-Glass Opacities in the Lung Cancer Spectrum. Ann Thorac Surg 2020;110:1796-804.
2. Cai Y, Chen T, Zhang S, et al. Correlation exploration among CT imaging, pathology and genotype of pulmonary ground-glass opacity. J Cell Mol Med 2023;27:2021-31.
3. Rami-Porta R, Nishimura KK, Giroux DJ, et al. The International Association for the Study of Lung Cancer Lung Cancer Staging Project: Proposals for Revision

of the TNM Stage Groups in the Forthcoming (Ninth) Edition of the TNM Classification for Lung Cancer. J Thorac Oncol 2024;19:1007-27.
4. Hu P, Liu M, Gu H, et al. Predicting the clinical prognosis of non-small cell lung cancer patients by predicting ALOX5 expression: a radiomics model. J Thorac Dis 2025;17:1387-99.
5. Tan N, Li Y, Ying J, et al. Histological transformation in lung adenocarcinoma: Insights of mechanisms and therapeutic windows. J Transl Int Med 2024;12:452-65.
6. Sidorenkov G, Stadhouders R, Jacobs C, et al. Multi-source data approach for personalized outcome prediction in lung cancer screening: update from the NELSON trial. Eur J Epidemiol 2023;38:445-54.
7. Sun F, Huang Y, Yang X, et al. Solid component ratio influences prognosis of GGO-featured IA stage invasive lung adenocarcinoma. Cancer Imaging 2020;20:87.
8. Palacios-Acedo AL, Mège D, Crescence L, et al. Platelets, Thrombo-Inflammation, and Cancer: Collaborating With the Enemy. Front Immunol 2019;10:1805.
9. Quach ME, Chen W, Li R. Mechanisms of platelet clearance and translation to improve platelet storage. Blood 2018;131:1512-21.
10. In 't Veld SGJG, Arkani M, Post E, et al. Detection and localization of early- and late-stage cancers using platelet RNA. Cancer Cell 2022;40:999-1009.e6.
11. Best MG, Sol N, Kooi I, et al. RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. Cancer Cell 2015;28:666-76.
12. Franco AT, Corken A, Ware J. Platelets at the interface of thrombosis, inflammation, and cancer. Blood 2015;126:582-8.
13. Goubran HA, Burnouf T, Stakiw J, et al. Platelet microparticle: a sensitive physiological "fine tuning" balancing factor in health and disease. Transfus Apher Sci 2015;52:12-8.
14. Carmeliet P, Jain RK. Angiogenesis in cancer and other diseases. Nature 2000;407:249-57.
15. Plantureux L, Mège D, Crescence L, et al. Impacts of Cancer on Platelet Production, Activation and Education and Mechanisms of Cancer-Associated Thrombosis. Cancers (Basel) 2018;10:441.
16. Plantureux L, Crescence L, Dignat-George F, et al. Effects of platelets on cancer progression. Thromb Res 2018;164 Suppl 1:S40-7.
17. Sromek M, Głogowski M, Chechlińska M, et al. Persistent and novel changes in plasma microRNA profiles in patients with non-small cell lung cancer following tumour resection. Transl Lung Cancer Res 2025;14:677-706.
18. Akbani R, Ng PK, Werner HM, et al. A pan-cancer proteomic perspective on The Cancer Genome Atlas. Nat Commun 2014;5:3887.
19. van der Meijden PEJ, Heemskerk JWM. Platelet biology and functions: new concepts and clinical perspectives. Nat Rev Cardiol 2019;16:166-79.
20. Mandel J, Casari M, Stepanyan M, et al. Beyond Hemostasis: Platelet Innate Immune Interactions and Thromboinflammation. Int J Mol Sci 2022;23:3868.
21. Schlesinger M. Role of platelets and platelet receptors in cancer metastasis. J Hematol Oncol 2018;11:125.
22. Tian B, Pang Y, Gao Y, et al. A pan-cancer analysis of the oncogenic role of Golgi transport 1B in human tumors. J Transl Int Med 2023;11:433-48.
23. Zhang Q, Bi Z, Song X, et al. Tumor-educated platelet SNORA58, SNORA68 and SNORD93 as novel diagnostic biomarkers for esophageal cancer. Future Oncol 2023;19:651-61.
24. Liefaard MC, Moore KS, Mulder L, et al. Tumour-educated platelets for breast cancer detection: biological and technical insights. Br J Cancer 2023;128:1572-81.
25. Li T, Guo T, Liu H, et al. Platelet derived growth factor BB mediates pancreatic cancer malignancy via regulation of the Hippo/Yes associated protein signaling pathway. Oncol Rep 2021;45:83-94.
26. Liu L, Lin F, Ma X, et al. Tumor-educated platelet as liquid biopsy in lung cancer patients. Crit Rev Oncol Hematol 2020;146:102863.
27. Sheng M, Dong Z, Xie Y. Identification of tumor-educated platelet biomarkers of non-small-cell lung cancer. Onco Targets Ther 2018;11:8143-51.
28. Wang Y, Hou K, Jin Y, et al. Lung adenocarcinoma-specific three-integrin signature contributes to poor outcomes by metastasis and immune escape pathways. J Transl Int Med 2021;9:249-63.
29. Hu Y, Zeng C, Li J, et al. TRIM27 revealing by tumor educated platelet RNA-sequencing, as a potential biomarker for malignant ground-glass opacities diagnosis mediates glycolysis of non-small cell lung cancer cells partially through HOXM1. Transl Lung Cancer Res 2024;13:2307-25.
30. Wen Z, Fu F, Zhao Y, et al. Residual tumor descriptors proposed by the International Association for the Study of Lung Cancer may not be applicable to stage I and ground-glass opacity-featured non-small cell lung cancer. Transl Lung Cancer Res 2023;12:2157-68.

31. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 2011;17:10-2.
32. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015;12:357-60.
33. Shen S, Park JW, Lu ZX, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A 2014;111:E5593-601.
34. Page C, Pitchford S. Platelets and allergic inflammation. Clin Exp Allergy 2014;44:901-13.
35. Labelle M, Begum S, Hynes RO. Direct signaling between platelets and cancer cells induces an epithelial-mesenchymal-like transition and promotes metastasis. Cancer Cell 2011;20:576-90.
36. Ding S, Dong X, Song X. Tumor educated platelet: the novel BioSource for cancer detection. Cancer Cell Int 2023;23:91.
37. Sciarrillo R, Wojtuszkiewicz A, Assaraf YG, et al. The role of alternative splicing in cancer: From oncogenesis to drug resistance. Drug Resist Updat 2020;53:100728.
38. Yu Z, Tang H, Chen S, et al. Exosomal LOC85009 inhibits docetaxel resistance in lung adenocarcinoma through regulating ATG5-induced autophagy. Drug Resist Updat 2023;67:100915.
39. Sharangdhar T, Sugimoto Y, Heraud-Farlow J, et al. A retained intron in the 3'-UTR of Calm3 mRNA mediates its Staufen2- and activity-dependent localization to neuronal dendrites. EMBO Rep 2017;18:1762-74.
40. Taylor R, Hamid F, Fielding T, et al. Prematurely terminated intron-retaining mRNAs invade axons in SFPQ null-driven neurodegeneration and are a hallmark of ALS. Nat Commun 2022;13:6994.
41. Cai Q, Kim M, Harada A, et al. Alpha-1 Antitrypsin Inhibits Tumorigenesis and Progression of Colitis-Associated Colon Cancer through Suppression of Inflammatory Neutrophil-Activated Serine Proteases and IGFBP-3 Proteolysis. Int J Mol Sci 2022;23:13737.
42. Cai Q, Dozmorov M, Oh Y. IGFBP-3/IGFBP-3 Receptor System as an Anti-Tumor and Anti-Metastatic Signaling in Cancer. Cells 2020;9:1261.
43. D'Addio F, Assi E, Maestroni A, et al. TMEM219 regulates the transcription factor expression and proliferation of beta cells. Front Endocrinol (Lausanne) 2024;15:1306127.
44. Vysotskiy M, Zhong X, Miller-Fleming TW, et al. Integration of genetic, transcriptomic, and clinical data provides insight into 16p11.2 and 22q11.2 CNV genes. Genome Med 2021;13:172.
45. D'Addio F, Montefusco L, Lunati ME, et al. Targeting a novel apoptotic pathway in human disease. Bioessays 2023;45:e2200231.
46. Dionisio-Vicuña MN, Gutiérrez-López TY, Adame-García SR, et al. VPS28, an ESCRT-I protein, regulates mitotic spindle organization via Gβγ, EG5 and TPX2. Biochim Biophys Acta Mol Cell Res 2018;1865:1012-22.
47. Zhu C, Xie Y, Li Q, et al. CPSF6-mediated XBP1 3'UTR shortening attenuates cisplatin-induced ER stress and elevates chemo-resistance in lung adenocarcinoma. Drug Resist Updat 2023;68:100933.
48. Shi W, Hu D, Xing Y, et al. Deciphering the Oncogenic Role of VPS28 Modulated by miR-491-5p in Breast Cancer Cells Using In Silico and Functional Analysis. Front Mol Biosci 2021;8:634183.
49. Mezzofanti E, Ignesti M, Hsu T, et al. Vps28 Is Involved in the Intracellular Trafficking of Awd, the Drosophila Homolog of NME1/2. Front Physiol 2019;10:983.
50. Liu F, Gao A, Zhang M, et al. Methylation of FAM110C is a synthetic lethal marker for ATR/CHK1 inhibitors in pancreatic cancer. J Transl Int Med 2024;12:274-87.
51. Jacinto S, Guerreiro P, de Oliveira RM, et al. MPV17 Mutations Are Associated With a Quiescent Energetic Metabolic Profile. Front Cell Neurosci 2021;15:641264.
52. Ge W, Wang Y, Quan M, et al. Activation of the PI3K/AKT signaling pathway by ARNTL2 enhances cellular glycolysis and sensitizes pancreatic adenocarcinoma to erlotinib. Mol Cancer 2024;23:48.
53. Sol N, In 't Veld SGJG, Vancura A, et al. Tumor-Educated Platelet RNA for the Detection and (Pseudo) progression Monitoring of Glioblastoma. Cell Rep Med 2020;1:100101.
54. D'Ambrosi S, Nilsson RJ, Wurdinger T. Platelets and tumor-associated RNA transfer. Blood 2021;137:3181-91.

Cite this article as: Shao M, Li W, Cao J, Wang L, Xie S, Hu Y, Feng G, Azari F, Bertolaccini L, Liu W, He B. Identification and validation of diagnostic alternative splicing events in tumor-educated platelets for non-small cell lung cancer in patients with ground-glass opacity: a multicenter study. Transl Lung Cancer Res 2025;14(4):1395-1407. doi: 10.21037/tlcr-2025-287