Multi-omics Research on the Heterogeneity and Immune Landscape of Lung Adenocarcinoma with Ground-glass opacity

Huiwen Miao  
The First Affiliated Hospital, Zhejiang University School of Medicine  

Xiao Teng  
The First Affiliated Hospital, Zhejiang University School of Medicine  

Xiaodong Teng  
The First Affiliated Hospital, Zhejiang University School of Medicine  

Shaozi Fu  
The First Affiliated Hospital, Zhejiang University School of Medicine  

Meihui Gao  
Sir Run Run Shaw Hospital, Zhejiang University School of Medicine  

Heyun Xu (✉ xuheyun@zju.edu.cn)  
The First Affiliated Hospital, Zhejiang University School of Medicine  

Research Article  

Keywords: lung adenocarcinoma, ground-glass opacity, proteomics, CXCL13, macrophage  

Posted Date: October 21st, 2022  

DOI: https://doi.org/10.21203/rs.3.rs-2155357/v1  

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background:
Lung adenocarcinoma with ground-glass opacity (GGO) has been detected increasingly and now accounts for most lung cancer patients. Lung adenocarcinoma with GGO contains a complex ecosystem. The mechanism of lung adenocarcinoma with GGO remains largely elusive. We use mass spectrometry proteomics combined with metabolomics to understand how these characteristics achieve a long-term functional balance and the trend of changes in tumor progression at the cellular functional level.

Methods:
We initiated a prospective cohort study to characterize lung adenocarcinoma with GGO components or without GGO components. Tumor and para-cancer tissue samples were collected. Multi-omics including transcriptomics proteomics and metabonomics were performed.

Results:
We found lung adenocarcinoma with GGO had a relatively slow proliferation tumor cells and stronger immune cell infiltration in proteomic and transcriptomic analysis. The immune cell markers expression, including CD47, CD68, CD81, CD86, C1Q, SPP1, CXCL13, ALOX5AP and HPGD was found overexpression in lung adenocarcinoma with GGO, which indicated more immune cell infiltration. In metabolomic analysis, GAPDH, ENO1 and LDHA were highly expressed in pure-solid lung adenocarcinoma, and GPD1 was highly expressed in lung adenocarcinoma with GGO. The combined transcriptome and proteome analysis revealed that proteins with consistent differences mainly included GAPDH, MKI67, AGER, and CRYM. KEGG pathway enrichment analysis showed that several aliphatic acyclic compounds expression were higher in lung adenocarcinoma with GGO.

Conclusion:
We describe a functional homeostasis in lung adenocarcinoma with GGO, which was constructed by relatively slow proliferation tumor cells and stronger immune cell infiltration. Overexpression of CXCL13 drives the infiltration of immune cells, which means the formation of anti-tumor tertiary lymphatic structures. The dysfunction of macrophage may be an important marker of this progression.

Introduction
Ground-glass opacity (GGO) of the lung refers to increased attenuation of the pulmonary vasculature without concealment. A previous follow-up study found that patients with early-stage lung adenocarcinoma with GGO had a lower recurrence rate and a better prognosis than patients with lung
adenocarcinoma without GGO under the same TNM Stage\textsuperscript{[1]}. This would suggest that lung adenocarcinoma with GGO possesses a more inert biological behavior. Therefore, the treatment strategy for this subtype of lung adenocarcinoma should not be overly aggressive\textsuperscript{[2,3]}.

Although a long-term follow-up strategy has been recommended for treating lung adenocarcinoma with GGO, a comprehensive understanding of the pathophysiology and mechanisms of this subtype of lung adenocarcinoma is conducive to a personalized and effective treatment plan. Despite extensive research into the mechanism of lung adenocarcinoma with GGO, it remains largely elusive. This is because lung tumors contain a complex ecosystem of malignant cells, immune cells, and stromal cells et al\textsuperscript{[4]}. The tumor develops under the immune system selection pressure. Furthermore, the immune pressure participates in shaping genetic heterogeneity levels\textsuperscript{[5]}. Therefore, tumor cell heterogeneity and the various types of the tumor microenvironment (TME) jointly shape the tumor behavior\textsuperscript{[6]}.

According to previous studies, the lung adenocarcinoma with GGO has a lower mutation load, fewer copy number variations, and a lower metabolism rate\textsuperscript{[7]}. Moreover, single-cell sequencing studies revealed that lung cancer with GGO had higher immune cell infiltrations\textsuperscript{[8,9]}. The heterogeneity and immune landscape of early (pT\textsubscript{1}N\textsubscript{0}M\textsubscript{0}, stage IA) stage of lung adenocarcinomas with GGO were studied at the cellular functional level using mass spectrometry proteomics combined with metabolomics to understand how these characteristics achieve a long-term functional balance and the trend of changes in tumor progression.

**Methods**

**Sample collection**

Tumor and para-cancer tissue samples were collected from 18–70 years old patients with lung adenocarcinoma less than 3 cm (cT1N0M0, Stage IA). Exclusion criteria: (1) postoperative pathological Stage higher than Stage IA; (2) multiple primary lung cancer; (3) history of the malignant tumor within last five years; (4) history of chemotherapy. The study was approved by the Ethics Committee of the First Affiliated Hospital of Medicine School of Zhejiang University (2021IIT fast review No.663) and obtained written informed consent from each patient.

**CT evaluation and grouping**

All patients underwent a high-resolution plain Computerized Tomography (CT) scan to evaluate the lesions before surgery. The nodules with GGO were classified as groups with ground-glass opacity (AD-G) (26 cases), while nodules without GGO were classified as a group of pure solid (AD-S) lung adenocarcinoma (24 cases). Para-cancer tissue (locating para-cancer based on the distance (2 cm) from cancer) was set as the control group (NC). We further classified AD-G into two groups: AD-G-H (GGO component more than 50%) and AD-G-L (GGO component less than 50%).

**Transcriptomics**
Total RNA was extracted using the Trizol method, and the purity and integrity of RNA were detected using NanoPhotometer and Agilent 2100 bioAnalyzer, respectively. The starting RNA for database construction was total RNA, and the total amount was $\geq 1 \mu g$. The database was constructed using Illumina's NEBNext® UltraTM RNA Library Prep Kit. Following approval on the inspection of the constructed database, Illumina sequencing was performed after pooling different databases based on the effective concentration and target offline data amount requirements. The original data must be filtered, which includes removing adapter reads, reads containing N (N indicates that basic group information cannot be determined), and low-quality reads (reads with Qphred $\leq 20$ accounting for more than 50% of the entire read length) to ensure the quality and reliability of data analysis. Meanwhile, Q20, Q30, and GC content were estimated from the clean data obtained in the study. All subsequent analyses were based on clean data. The reference genome and gene model annotation files were downloaded directly from the genome website. A reference genome index was constructed using HISAT2 V2.0.5, and clean reads of the paired terminal were compared with the reference genome using HISAT2 V2.0.5. StringTie (1.3.3b) was used to predict new genes, and FeatureCounts (1.5.0-P3) to calculate the readings mapped to each gene. Then, Fragments Per Kilobase of transcript per Million mapped reads (FPKM) was calculated for each gene based on its length, and the reading mapped to each gene was also calculated. DESeq2 software (1.16.1) was applied for differential expression analysis between the two comparison groups. ClusterProfiler (3.4.4) software was devoted to the enrichment analysis of differentially expressed genes. PPI analysis of differentially expressed genes was performed according to the STRING database.

Proteomics

The samples were treated with Homogenate + SDT lysis and the supernatant was collected for protein quantification using the protein BCA kit. Samples were analyzed by SDS-PAGE electrophoresis. FASP enzymolysis was performed after extracting the same amount of protein from each sample, and the peptide was desalted and quantified. The peptide samples were combined to make a mixture. The peptide samples were graded on an Agilent 1260 Infinity II HPLC system and merged into six components for data-dependent acquisition (DDA) analysis. A spectrogram database was established for data-independent acquisition (DIA) analysis. Before DIA analysis, the peptide of each sample was diluted to 10 ng/µL with 0.1% FA and mixed with the iRT peptide. Before sample loading, 200 ng peptide was obtained from each sample and desalted with Evotips. After Evosep One chromatographic separation, PASEF mode of the timsTOF Pro (Bruker, Bremen, Germany) mass spectrometer was utilized for mass spectrometry analysis. Regarding DDA analysis, the ion accumulation or release time was set to 100 ms, the ion mobility was scanned from 0.85 to 1.3 Vs/cm², and mass spectrograms in the m/z 100–700 range were scanned under positive electrospray mode. A complete TIMS-MS acquisition time was 0.53 s, including four PASEF MS/MS scans. The quadrupole isolation window was defined as a function of TIMS scan time during DIA analysis. Every 100 ms of diaPASEF scanning, two windows with a size of 25 m/z were established. The DIA raw data were processed and analyzed by Spectronaut (Biognosys AG, Switzerland) with the retention time prediction type set to dynamic iRT and default settings. Spectronaut's Pulsar database search engine was used to generate a spectrographic database from the DDA-PASEf raw
data, and the FDR peptide level was controlled at 1%. Trypsin enzymolysis was performed to allow two missing sites. The fixed modifier of database search parameters was Carbamidomethyl (C), and variable modifiers were Oxidation (M) and acetyl (Protein n-term) N-terminal acetylation. After passing the t-test, proteins are considered differentially expressed if their PG.Q-value < 0.05 and their multiple changes in up-regulation or down-regulation were > 1.5. The differential proteins were compared and annotated using Blast2GO and KOALA (KEGG Orthology and Links Annotation). Moreover, GO and KEGG enrichment analysis were also conducted.

**Metabonomics**

Samples were slowly thawed at 4 °C. After ultrasonic centrifugation, 800 μL methanol/acetonitrile (1:1, V/V) was added to the homogeneous solution for metabolite extraction, and the supernatant was collected for drying. Before LC-MS analysis, the samples were re-dissolved in 100 μL acetonitrile/water (1:1, V/V) solvent, centrifuged at 14000 g at 4 °C for 15 min, and the supernatant was analyzed. HILIC column of Agilent 1290 Infinity LC ULTRA High-Performance Liquid Chromatography (UHPLC) was used for separation. AB Triple TOF 6600 high-resolution mass spectrometer was used for primary and secondary data acquisition. UHPLC-Q-TOF-MS was used to detect metabolites in the samples. The liquid mobile phase in the experiment consisted of A, water +25 mM ammonium acetate +25 mM ammonia, water, and B, acetonitrile. The elution time is 12 min. The mass spectrometry detection included positive and negative modes, with a voltage of ± 5500 V, a scanning m/z range of 60–1000 Da, collision energy of 35 ± 15 eV, and a scanning speed of 0.05 s/Spectra. The original data were converted into mzXML format by ProteoWizard. Then, XCMS software was used for peak alignment, retention time correction, and peak area extraction. The metabolite ID was confirmed by the retention time, molecular mass (molecular mass error < 10 PPM), secondary fragmentation spectrum, collision energy, and other information in the local database. Data analysis included univariate statistical analysis, multidimensional statistical analysis, differential metabolite screening, differential metabolite correlation analysis, and KEGG pathway analysis. Metabolites meeting the t-test PG.Q-value < 0.05 and OPLS-DA VIP > 1 were regarded as differentially expressed metabolites. Subsequently, a KEGG pathway analysis was performed.

**Immunohistochemical Study**

Immunohistochemical (IHC) analysis was performed by a fully automated system (Benchmark XT System; Ventana Medical Systems, Tucson, AZ, USA) at our institution. IHC was performed by following antibodies: an anti-BCA1 antibody (ab246518, 1:1000, Abcam), anti-CD81 antibody (200861-T08, 1:1000, Sino Biological), anti-C1QA antibody (ab189922, 1:1000, Abcam), anti-CD47 antibody (12283-T26, 1:2000, Sino Biological), anti-CD68 antibody (11192-MM01, 1:1000, Sino Biological), anti-CD63 antibody (ab271286, 1:200, Abcam), anti-RAGE antibody (11629-T24, 1:2000, Sino Biological), anti-CRYM antibody (105240-T40, 1:1000, Sino Biological), anti-CLDN3 antibody (100946-T08, 1:500, Sino Biological). Appropriate positive and negative controls were used. The results were reviewed by an experienced pathologist.
**Statistical analysis**

Experimental results were expressed as mean ± standard deviation (SD). A t-test was used to compare differences between groups, and Q-value < 0.05 was considered statistically significant. SPSS 22.0 was used for all statistical analyses. The original data was converted into MzXML format by ProteoWizard, and then XCMS software was used for peak alignment, retention time correction, and peak area extraction. First, metabolite structure identification and data pre-processing were carried out for XCMS extracted data, then the quality evaluation of experimental data was made, and finally, the data analysis. Data analysis included univariate statistical analysis, multidimensional statistical analysis, differential metabolite screening, correlation analysis of differential metabolite, KEGG pathway analysis, and other contents.

**Results**

**Metabonomics results and conjoint analysis**

The transcriptome difference between lung adenocarcinoma with ground-glass opacity (AD-G) (26 cases) and pure solid lung adenocarcinoma (AD-S) (24 cases) was less than the proteome (Fig. 1A, 1B and 1C). The combined transcriptome and proteome analysis (Fig. 1D) revealed consistent and opposite differences. The proteins with consistent differences mainly included GAPDH, MKI67, AGER, and CRYM. The KEGG pathway enrichment results are displayed in Fig. 1E. However, SNRPC contained proteins with opposite differences (Fig. 1F).

KEGG pathway enrichment analysis of proteome was performed between AD-G and AD-S (Fig. 1G). Positive and negative ion patterns were significantly different in metabolite expression among the three groups (Fig. 1H and 1I). Several aliphatic acyclic compounds expression was the highest in AD-G, compared with AD-S and para-cancer tissue, including histamine, 15-Deoxy-delta-12,14-PGJ2, N6-methyl-L-Lysine, Glutaraldehyde, Deoxycarnitine, L-Carnitine, Phenylacetaldehyde and Aspartic acid. The combined proteomic and metabolomic analysis suggested that the proteomic difference between AD-G and AD-S was greater (Fig. 1J).

**Immune Cell-related Markers**

The immune cell markers expression, including CD21, CD68, CD3D, CD3G, IGHV3-35, IGHV3-15, IGHV1-46, and chemokine CXCL13, was highest in AD-G, suggesting more immune cell infiltration (Fig. 2A). ALOX5AP and HPGDS, associated with macrophages, were highly expressed in the lung adenocarcinoma with GGO (Fig. 2B). We discovered that only CD68 expression was higher in the GGO with less than 50% solid component than in the GGO with more than 50% solid component (Fig. 2C). We also found that lung adenocarcinoma with GGO had the lowest CD86 expression (Fig. 2C).
We also found that CD 68 and SPP1, CD68 and C1QA were co-expression on the macrophage in the IHC double-staining. (Fig. 2D)

**Hla Difference**

HLA-DRA and HLA-DRB5 expressions were highest in AD-G (Fig. 3A). In the subgroup analysis, HLA-DRB1 and HLA-DRB5 were lower in AD-G-L compared with AD-G-H (Fig. 3B). These major histocompatibility-II complex (MHC-II complex) was found on specialized APC but can also be expressed in some tumor cells. MHC-II was critical for antigenic presentation of CD4 + T lymphocytes and associated with a better prognosis\(^\text{[11]}\).

**Tumor Cell Protein**

The tumor suppressor gene expressions were different in the three groups. The ZFPL1, SPP1, and CRABP2 expressions were the highest in AD-S, and POLQ, ITLN1, and APOF expressions were the highest in para-cancer tissues (Fig. 4A). However, these gene expressions in AD-G were between para-cancer tissues and AD-S. Cyclin CDK1, CDK2, CDK4, CDK7, MKI67, and CD274 expression levels were the highest in AD-S, and there was no significant difference in cyclin expression between AD-G and para-cancer tissues (Fig. 4B). CD81, CRYM, CLDN3, CD47, TMSB10, and other proteins expression were highest in AD-G (Fig. 4C). As the mechanism of some proteins mentioned above has not been reported, immunohistochemical staining was used to locate CLDN3 CRYM, CD81, CD47, and TMSB10. We found CLDN3, and CRYM expressed in tumor cytoplasm and CD81 in the tumor cell membrane. (Fig. 4D) CD47 was expressed in normal epithelium and tumor cell membrane of GGO. (Fig. 4D) TMSB10 staining was located in the matrix. (Fig. 4D)

**Metabolism-related Proteins And Metabolomics Results**

Glycolysis and pyruvate oxidation related glyceraldehyde 3-phosphate dehydrogenase (GAPDH), enolase 1 (ENO1), and lactate dehydrogenase (LDHA) were highly expressed in pure-solid lung adenocarcinoma, and glycerol 3-phosphate dehydrogenase (GPD1) was highly expressed in AD-G. (Fig. 5A) Three proteins involved in mitochondrial transport and fatty acid oxidation, carnitine palmitoyl transferase 1A (CPT1A), electron transfer flavin protein subunit α (ETFA), and hydroxy-CoA dehydrogenase subunit α demonstrated no statistical significance among the three groups. This was consistent with the clinical manifestations. Pure-solid lung adenocarcinoma preferred to use glycolysis, an inefficient metabolic pathway for energy metabolism. However, AD-G had lower energy requirements than AD-S, resulting in lower glycolysis-associated gene expression\(^\text{[12]}\).

**Discussion**
With the progressive rise of solid components of AD-G models were designed for studying the early progression of lung adenocarcinoma. AD-G showed a relatively inert biological activity even after progression to invasive lung adenocarcinoma, suggesting that it may have a unique tumor cell and tumor microenvironment composition.

AD-G was an intermediate between normal tissue and lung adenocarcinoma with pure solid, with various oncogenes and tumor suppressor gene expression also falling between the two. Ki67 and various cyclin expression levels suggest that AD-G have not entered the proliferation state. Some AD-G specific proteins may be important regulators of tumor progression. Through IHC staining localization and literature analysis, we found that CLDN3 may be the main "intrinsic cause" of tumor growth inhibition. CLDN3 has preserved the epithelial phenotype and suppressed tumor growth in various cancers, including lung cancer[^13,^14].

Additionally, CRYM, a novel antagonist of T3 that participates in androgen-mediated signal transduction in prostate cancer, might act as a regulator to prevent prostate cancer positive growth[^15]. However, the mechanisms by which AD-G specific proteins regulate tumor progression are unknown. Hence, further research is warranted to confirm whether CRYM and other AD-G specific proteins could modulate lung adenocarcinoma progression.

Cell markers such as CD68, CD3D, CD3G, CD21, IGHV3-35, IGHV3-15, and IGHV1-46 expression indicate an increasing trend in macrophages, T, and B lymphocytes recruitment in AD-G tissues, consistent with single-cell sequencing results[^8,^9].

CXCL-13 plays a fundamental role in the formation of tumor-associated tertiary lymphoid structures[^16]. Interestingly, CXCL-13 expression was significant higher in AD-G compared with para-cancer tissue and AD-S. However, IHC studies showed that macrophage in AD-G did not express CXCL-13. Furthermore, a B cell activator, CD 86 expression was lowest in AD-G, compared with AD-S and para-cancer tissue[^17]. It indicted that macrophage did not participate the formation of tertiary lymphatic structures. Hence, it maybe hematopoietic lymph tissue inducer cells or other alternative cells (e.g. Th17 cells producing IL-17 through RORyt transcription factor and CD8 + T cells in chronic inflammatory microenvironment) in AD-G drove the formation of tertiary lymphatic structures[^18].

Moreover, our metabolomics findings revealed an increase in specificity of histamine, 15-deoxy-Δ12,14-prostaglandin J2 (15-PGJ2), and aliphatic acyclic compounds. Several studies have confirmed that 15-PGJ2 had significant anticancer effects[^19–21]. It also has anti-inflammatory, antiangiogenic, and pro-apoptotic properties. Moreover, Sonja demonstrated that macrophages could be one of the major sources of prostaglandin D2 during inflammatory responses[^18]. In this study, AD-G showed a higher macrophage infiltration expression. Additionally, prostaglandin D2 is a precursor of 15d-PGJ2. Hence, macrophage-induced 15-PGJ2 expression is believed to play a role in maintaining a relatively stable immune microenvironment in AD-G.
The MHC-II molecules such as HLA-DRA and HLA-DRB5 expression were higher in AD-G, indicating a better antigen presentation ability since macrophages\cite{22}, dendritic cells, and CD86 + antigen-presenting B cells express these proteins. However, CD86 expression was lowest in AD-G compared with para-cancer and AD-S. It may indicate that antigen-presenting activity of CD86 + cells may not play an essential role of antigen presentation in AD-G\cite{23}. We also found the high CD47 expression in AD-G, which may indicate that tumors may evade macrophage phagocytosis via the CD47/SIRP\alpha axis\cite{24,25}. Furthermore, polarization pattern of macrophage may distinguish its function in tumor cells. Previous studies on the polarization type of AD-G yielded conflicting findings\cite{8,9}. In our study, ALOX5AP and HPGDS were highly expressed in AD-G, indicating macrophages underwent M1-type polarization and induced the histamine release from mast cells via pro-inflammatory lipid mediators\cite{26,27}. They also promoted PGD2 production, a precursor of anti-inflammatory prostaglandin 15-deoxy-delta-12,14-PGJ2\cite{28}. In general, macrophages in AD-G showed a M1-polarization tendency, and may be a key factor of initiating lymphoid tissue formation\cite{29}. However, we also observed that C1Q + and SPP1 + macrophages in AD-G. C1Q + macrophage was thought to promote lung cancer adhesion, migration and proliferation\cite{30}. SPP1 + macrophage was associated with lung cancer activation and early lymph node metastasis\cite{31,32}. It indicated the complexity and diversity of macrophages. The function of these cells remains to be further investigated\cite{33}.

As mentioned previously, AD-G had greater lymphocyte infiltration, a relatively low tumor cell metabolism, and a low proliferation index. Consequently, another special mechanism may regulate the immune system pressure with GGO. Proteomic and immunohistochemistry analyses revealed that the CD81 expressions, commonly used as markers for exosomes and other tiny extracellular vesicles (EVs)\cite{34}, were higher in AD-G compared to AD-S, suggesting that AD-G might have higher EV secretion rate. Exosomes played an important role in intercellular communication\cite{34,35}. The relationship between tumor-associated EVs and the immune system is complex\cite{36,37}. EVs can promote both pro-tumor and anti-tumor immunity, suggesting that the EVs functions are mostly dependent on cell type and EV cargo. EVs can transfer tumor antigens to APC and present them to T cells to induce primary cytotoxic immune response. However, it can also mediate immunosuppression, which is important for tumor progression. CD81 was essential for the immunosuppressive function in both regulatory T cells and congenital myeloid suppressor cells\cite{38}. Furthermore, high expression of CD47 made it possible for tumor-associated EVs related to GGO cancer cells to avoid phagocytosis and have a longer survival time\cite{24,39}. Therefore, tumor associated EVs in AD-G may avoid phagocytosis and have a longer survival time. In summary, immune escape of AD-G may be related to tumor associated EVs though high CD81 and CD47 expression, further research on the tumor EV contents (e.g., miRNA) will be important in understanding the immunological homeostasis of GGO.

Mitochondria are key centers for regulating metabolism and cell signal transduction and play a role in tumor progression and treatment resistance\cite{40}. Our study discovered only glucose-related alterations in protein expression. Moreover, the GAPDH and GPD1 modifications were consistent with prior research.
findings, and enhanced glucose uptake is a functional biomarker with worse overall survival (OS) and disease-free survival (DFS). In contrast, increased tumor GPD1 levels were correlated with better OS and DFS\textsuperscript{[12]}. This indicates mitochondria are more important in solid lung adenocarcinoma. Additionally, mitochondria may be a therapeutic target in pure solid lung adenocarcinoma.

Our study found that U1 small nuclear ribonucleoprotein C (SNRPC) expression was higher in lung adenocarcinoma with GGO compared with pure solid lung adenocarcinoma. As a part of U1 small nuclear ribonucleoprotein (U1 snRNP), SNRPC is crucial to initiating and regulating precursor messenger RNA splicing\textsuperscript{[41]}. Existing studies found that U1 snRNP plays an important role in the cancer cells migration and invasion in-vitro\textsuperscript{[42]}. Since the combination and functional activity of SNRPC and U1 snRNA are independent of all subunits of U1 snRNP\textsuperscript{[41,43]}. SNRPC regulates the protein metabolism and liver cancer immune process. SNRPC in the AD-G might be up-regulated in post-transcription modification. Particulate matter (PM)\textsubscript{2.5} in the air can induce SNRPC upregulation in human lung cancer cells and pro-inflammatory cytokines secretion\textsuperscript{[44]}. Therefore, exposure to (PM)\textsubscript{2.5} and up-regulated SNRPC might play an important role in lung adenocarcinoma with GGO.

Nevertheless, this study still has several shortcomings. Firstly, the proteomics results are the overall results of tissue detection, the results are relatively macro and precise cell functions cannot be analyzed. Secondly, the solid components gradually formed in AD-G tissue is considered as tumor progression, and the spatial heterogeneity of AD-G needs to be further studied. In future, more in-depth research of rapid formation of tertiary lymphoid structures in the early stage of tumorigenesis and regulation of TME is required.

**Conclusions**

We describe a functional homeostasis in lung adenocarcinoma with GGO, which was constructed by relatively slow proliferation tumor cells and stronger immune cell infiltration. Overexpression of CXCL13 drives the infiltration of immune cells, which indicated the formation of anti-tumor tertiary lymphatic structures. We also found that tumor cells escape immune surveillance through cell surface proteins and exosomes. The dysfunction of macrophage may be an important marker of this progression. However, it is unknown what causes the higher expression of CXCL13 and the formation of a perfect anti-tumor tertiary lymphatic structure in lung adenocarcinoma with GGO.

**Abbreviations**
Declarations

Ethical Approval

This study was approved by the Ethics Committee of the First Affiliated Hospital of Medicine School of Zhejiang University (2021IIT fast review No.663)

Author Contributions:

H.M. and X.T. conceived, performed the bioinformatics analysis and wrote the manuscript. S.F. and M.G. collected the tumor samples and analyzed the data of clinical characteristics. X.T. reviewed the immunohistological analysis result. H.X. supervised this study.

Competing Interests statement:

All authors have no conflicts of interest associated with this study.

Availability of data and materials

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

Funding

This study had no funding support.

Authors’ contributions

HX designed this study. HM, XT, XT and SF performed the majority of the experiments. MG contributed to statistical analysis. HM and XT wrote the manuscript. HX revised the manuscript critically. All the authors read and approved the final version of the manuscript.

Consent for publication

All the authors read and approved the final version of the manuscript to be published.
Competing interests

The authors declare that they have no conflicts of interest.

References


Figures
Figure 1

A The number of differential gene expression display among three groups (AD-G vs. AD-S vs. NC); B The number of differential protein expression among three groups (AD-G vs. AD-S vs. NC); C Venn diagram of differential protein and transcript expression between AD-G and AD-S; D The combined transcriptome and proteome analysis of consistent and opposite differences between AD-G and AD-S; E The KEGG pathway enrichment results of consistent difference in combined transcriptome and proteome analyses between AD-G and AD-S; F The difference expression of SNRPC among three groups (AD-G vs. AD-S vs. NC); G The KEGG pathway enrichment results of proteome between AD-G and AD-S; H Positive ion patterns were significantly different in metabolite expression among the three groups (AD-G vs. AD-S vs. NC); I Negative ion patterns were significantly different in metabolite expression among the three groups (AD-G vs. AD-S vs. NC); J Combined proteomic and metabolomic analysis between AD-G and AD-S

Figure 2

A Immune cell markers expression: CD21, CD68, CD3D, CD3G, IGHV3-35, IGHV3-15, IGHV1-46, and CXCL13, was highest in AD-G; B ALOX5AP and HPGDS were highly expressed in AD-G; C CD68 expression was higher in the AD-G with less than 50% solid component than more than 50% solid component. CD86
expression was lowest in AD-G; D SPP1 and CD 68, C1QA and CD68 were co-expression on the macrophage in the IHC double-staining.

Figure 3

A HLA-DRA and HLA-DRB5 expressions were highest in AD-G; B HLA-DRB1 and HLA-DRB5 were lower in AD-G with less than 50% solid component than more than 50% solid component.
Figure 4

A The ZFPL1, SPP1, and CRABP2 expressions were the highest in AD-S, while POLQ, ITLN1, and APOF expressions were the highest in NC; B Cyclin CDK1, CDK2, CDK4, CDK7, MKI67, and CD274 expression levels were the highest in AD-S, and there was no significant difference in cyclin expression between AD-G and NC; C CD81, CRYM, CLDN3, CD47 and TMSB10 expression were highest in AD-G; D CLDN3, and CRYM expressed in tumor cytoplasm and CD81 in the tumor cell membrane. CD47 was expressed in normal epithelium and tumor cell membrane of AD-G.
Figure 5

A GAPDH, ENO1 and LDHA were highly expressed in AD-S compared with AD-G. However, GPD1 was highly expressed in AD-G.