ORIGINAL RESEARCH

# Identification of Genes and Key Pathways Associated with the Pathophysiology of Lung Cancer and Atrial Fibrillation

Dongmei Yang, MD; Yong Chen, MD; Yonglin Yu, BD; Xiaoju Chen, PhD

## ABSTRACT

**Background** • Lung cancer is a malignant tumor originating from respiratory epithelial cells in the bronchi, bronchioles, and alveoli, often associated with atrial fibrillation; However, there is a lack of in-depth understanding of its genetic basis and molecular mechanisms. Our goal is to study the genes and signaling networks associated with cancer and atrial fibrillation.

**Materials and methods** • We obtained microarray datasets for lung tumors from the Gene Expression Omnibus (GEO) database and AF for this investigation: GSE30219, GSE79768, and screened the candidate specimens in both microarrays for differential genes at $P < .05$ using GEO2R. The outcomes were also examined using the Gene Ontology (GO) functional enrichment and Kyoto Encyclopedia of Gene and Gene Combinations (KEGG) pathway analysis methods. Using STRING and Cytoscape, protein interaction networks (PPI) were analyzed and visualized. The Molecular Complex Detection (MOCDE) plugin is responsible for filtering important compounds. Candidate genes are then screened by the cytoHubba plugin according to MCC criteria. After taking the intersection of the Hub genes by the Wayne diagram, the ROC curves were plotted separately by combining the data from one lung cancer dataset GSE19804, two AF datasets GSE41177/GSE14975 in the GEO database.

**Results** • An aggregate of 49 co-expressed differentially expressed genes (co-DEGs) were discovered in lung cancer/AF and healthy controls. Most co-DEGs were found in neutrophil activation, where they were linked to immunological response and interactions between cytokines and cytokine receptors, according to GO and KEGG pathway analyses. Furthermore, due of their significant connectedness in both the lung carcinoma and AF datasets, we chose six key genes. They are MNDA, HP, LYZ, S100A9, S100A8, and S100A12, among others.

**Conclusions** • The findings of this research indicate that the onset of lung cancer and AF depends on a small number of distinctive genes. We investigated the functional enrichment, differential gene expression, and PPI of DEGs in lung cancer and AF, and the results offer fresh perspectives on the discovery of prospective biomarkers and priceless therapeutic precursors in these two diseases. (*Altern Ther Health Med.* [E-pub ahead of print.])

**Dongmei Yang**, MD; **Yong Chen**, MD, Department of Respiratory and Critical Care Medicine, The Affiliated Hospital of North Sichuan Medical College, Nanchong, Sichuan, China. **Yonglin Yu**, BD, Department of stomatology, The Affiliated Hospital of North Sichuan Medical College, Nanchong, Sichuan, China. **Xiaoju Chen**, PhD, Clinical Medical College & Affiliated Hospital of Chengdu University, Chengdu, Sichuan, China.

*Corresponding author: Xiaoju Chen,* PhD
*E-mail: 13989185102@sina.cn*

## INTRODUCTION

Lung cancer, also known as primary bronchial cancer of the lungs, is one of the malignant tumors with the highest morbidity and fatality rates in China as well as the rest of the globe. According to data from international cancer centers, most early-stage lung cancers have atypical clinical manifestations, and most patients are in the middle and late stage when they present with symptoms.[1]

Atrial fibrillation (AF) is the most frequent arrhythmia managed in clinical practice, and it is linked to an increased risk of death, stroke, and peripheral embolism. A GBD (Global Burden of Disease) investigation[2] showed that both the incidence and prevalence of AF as well as the resulting mortality and morbidity have been expanding globally, with a difference of roughly four times in the prevalence of AF around the world, with the greatest incidence in North America as well as in some European regions. As of 2019, there are approximately 59.7 million cases of atrial fibrillation, including atrial flutter. China, the Middle East, South Asia, Southeast Asia, and Africa have the lowest rates of AF. Old age, obesity, diabetes, heart disease, pulmonary embolism (PE), pneumonia, and cancer are risk factors linked to an increased risk of AF.[3] It is also linked to a raised risk of death from stroke, heart failure due to myocardial infarction, chronic kidney disease (CKD), and thromboembolism.[4]

The risk of getting AF varies depending on the kind of cancer; lung cancer has the highest risk[5] according to a recent Danish study on the incidence of AF in cancer sufferers. Furthermore, a nationwide population-based research from Korea found that among the primary forms of solid malignancies, lung cancer was associated with the highest prevalence of AF in the > 50-year-old age group,[6] which suggests that lung cancer and AF are co-morbidities. According to a research presented at the European Society of Cardiology (ESC) 2020 Congress, cancer patients have a greater incidence of atrial fibrillation (AF),[7] with 2.4% of cancer patients suffering from AF.[8] Furthermore, compared to persons without the condition, those with cancer may have a higher chance of getting AF.

Additionally, several studies have discovered that new-onset of AF in those with cancer may indicate more advanced cancer stages and a worse tumor prognosis.[9, 10] Some studies[9, 11, 12] have found that the main possible mechanisms by which AF often occurs in cancer patients are, (1) systemic inflammation that can promote atrial remodeling; (2) a combination of pain, emotional, and/or physically induced autonomic nervous system imbalance; (3) development of carcinoma in the heart tissue as well as surrounding tissues in episodic AF, either directly or by metastasis; (4) Interference of comprehensive cancer therapies such as surgery, chemotherapy, radiotherapy, and targeted therapies with the cardiovascular system's own homeostasis, leading to cardiomyocyte damage while causing reactions and increased oxidative stress to induce AF; (5) compression of nerves by tumor and secretion of multiple hormonal substances by tumor cells; and (6) other factors such as fluid imbalance, infections, abnormalities of the metabolism and electrolytes, cancer-related cardiac thrombus invasions that cause varying degrees of hemodynamic disturbances, and fluid imbalance during chemotherapy. In addition, AF may be an early sign of an occult tumor or an initial manifestation of systemic progression during treatment follow-up. The existence of risk factors in both groups may be one of the reasons of uncertainty about the underlying mechanism for the elevated long-term risk of tumors in AF patients. Inflammation, oxidative stress or a combination of factors may trigger and drive both the conditions.[13]

Thus, AF is found to be a common co-morbidity in cancer patients, and the factors predicting the prognosis of this high-risk group are not known. In this study, gene microarrays allowed the detection of changes in gene expression in patients with lung cancer combined with atrial fibrillation. The molecular processes and pathways of signaling identified may also assist in understanding the association of lung cancer and atrial fibrillation. The key genes identified could generate new research interests.

## MATERIALS AND METHODS

### Data sources

High-throughput sequencing and microarray datasets from research groups worldwide are submitted to the public Gene Expression Omnibus database (GEO) (http://www.ncbi.nlm.nih.gov/geo). We used the terms lung cancer and atrial fibrillation to look for relevant gene expression datasets. Two distinct expression profiles from two different sequencing platforms were chosen as the inclusion criterion, along with a maximum sample size. Additionally, the test samples that are included must be human. Finally, the two microarray datasets, GSE30219 and GSE79768 were downloaded (GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array; GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array). The GSE30219 dataset comprises 293 lung tumor specimens and 14 individuals with non-tumor lung specimens. Whereas the GSE115574 dataset comprises 31 healthy participants with sinus rhythm and 28 atrial fibrillation sufferers. R software (version 3.6.3) was used for data preparation, processing, and analysis.

### Identification of DEGs (differentially expressed genes) in lung cancer and atrial fibrillation

An effective bioinformatics-based analytic technique is the limma R package.[14] It was employed to locate the associated DEGs in lung cancer versus non-tumor patients and in sinus rhythm versus AF patients. Selection criteria for both lung cancer and AF datasets were set to $P < .05$ & $|\log2 \text{FC}| > 1$. The use of the above-mentioned data screening specifications helped us to distinguish between the two groups of DEGs. Thereafter, in the online analysis tool xiantaoxueshu (https://www.xiantao.love/), we used the DEGs as input for both the disorders to obtain the intersecting genes, which were displayed as Venn diagrams. We then used these crossed common genes for further analysis.

### Common DEGs' functional enrichment analysis

Gene Ontology (GO) analysis[15] and Kyoto Encyclopaedia of Genes & Genomes (KEGG) pathway enrichment analysis[16] were performed on crossed common DEGs, which include cellular components, biological processes, and molecular activities. A $P$ value of 0.05 or below was considered statistically significant.

### Hub gene identification, MCODE cluster modules, and PPI network analysis

The PPI network was examined using the Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org).[17] To identify the pathways driving lung cancer, total scores and functional relationships among proteins were examined. Utilizing the Cytoscape plug-in Molecular Complex Detection (MCODE), an open-source bioinformatics software platform, molecular interaction networks are visualized for analysis. The selection criteria we applied were as follows: Degree-cut-off = 2, Node Score Cut-off = 0.2, Maximum Depth = 100, and K-score = 2 are the

MCODE settings for filtering pertinent modules in PPI networks (parameters are set to default).[18]

### Choosing and evaluating hub genes

Top 10 genes were chosen using the MCC algorithm and the Cytoscape cytoHubba plug-in. The intersection set was then taken from the analysis of the linked core genes using MCODE, and it was depicted as a Venn diagram. Finally, the hub genes were examined using the GeneMANIA web database (https://genemania.org/).[19]

### Analysis of the hub genes using ROC

The subject operating characteristic (ROC) curve was constructed using the R package, and the area under the curve (AUC) of the hub genes was calculated by comparing the AUC as a model indicator. These findings demonstrated the genes' ability to provide accurate diagnoses. The endpoint was shown in figure 7. $P < .05$ was deemed to be statistically noteworthy.

### Statistical analyses

Statistical analyses were performed using the statistical Package for Social Sciences Version 19.0 (SPSS19.0). Student's *t* test and one-way ANOVA were used to compare the means of 2 groups or more. Mann-Whitney U-test were used to analyze the data of the associations. $P < .05$ was considered as significant difference.

## RESULT

### DEGs' differential expression in atrial fibrillation and lung cancer

When we compared 293 lung tumor specimens with 14 non-tumor lung specimen samples, and 14 AF control samples, 823 DEGs were removed from the lung cancer dataset GSE30219. After homogenization, 2541 and 131 DEGs were produced by adjusted-$P < .05$ treatment, correspondingly. These DEGs are shown in various colors in the volcano plot (Figure 1 and 2). The heat map is used to show the patterns of gene expression in lung cancer and AF, as well as the expression of these differentially expressed genes in each sample (Figures 1 and 2). Using Venn diagrams, 49 cross-shared genes were obtained for both diseases, as shown in Figure 3.

### Analysis of enriched KEGG and GO pathways

Functional enrichment and pathway analysis of 49 common genes in lung tumors and AF were carried out. The biological process (BP) largely involves changes in the activity of the toll-like receptor pathway, the cargo receptor pathway is affected, along with substantial changes in the activity of the receptors (rage receptor binding, cargo receptor activity, toll-like receptor, and scavenger receptor activity). The enrichment of cell outer membranes, including particular granule lumen and vesicle lumen, was a prominent focus of the changes in the cellular component (CC). Additionally, the molecular function (MF) shown in Figure 4 (A and B) focused mostly on inflammatory responses. Additionally, in the molecular function (MF) section (Figures 4A and 4B), the main immune responses considered were humoral immune response, neutrophil activation involved in immune response, antimicrobial humoral response, neutrophil degranulation, and inflammatory reactions involved in immune response.

A dot plot in (Figures 4C and 4D) shows how modifications in the KEGG pathway were most prevalent in numerous system disorders (legionellosis, malaria, fat digestion, and absorption), as well as in immune-related processes (IL-17 signaling pathway and cytokine-cytokine receptor interaction).

**Figure 1.** (A) Volcano Map of the GSE79768 Gene Chip. (B) Box-and-Whisker Evaluation Plot of the GSE79768 Gene Chip. (C) Identification of DEGs Heatmap of Genes in GSE79768 Gene Chip.

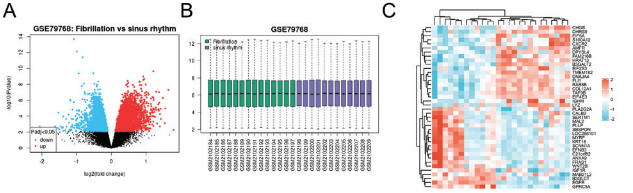


**Figure 2.** (A) Box-and-Whisker Evaluation Plot of the GSE30219 Gene Chip. (B) Volcano Map of the GSE30219 Gene Chip. (C) Identification of DEGs Heatmap of Genes in GSE30219 Gene Chip.

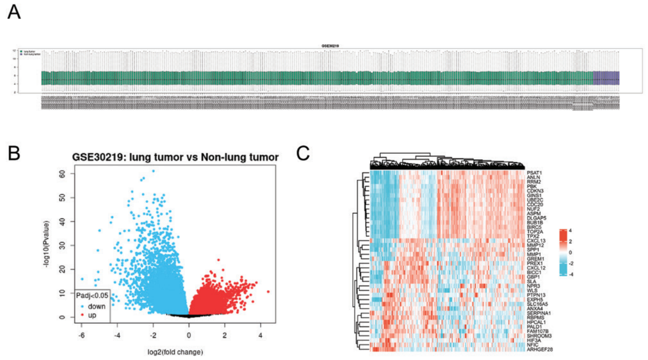


**Figure 3.** (A) Venn Diagram Depicting Lung Cancer/AF Crossover Genes

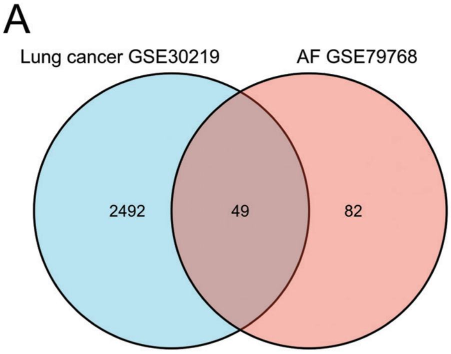


**Figure 4.** Common Genes Between Lung Cancer and AF were Analyzed for Function and Pathway Enrichment. Significant Gene Ontology Analysis of Biological Processes (BP), Cellular Components (CC), and Molecular Functions (MF) is Shown in (A–B). (C–D) Common Gene Enrichment Analysis Using KEGG Pathways.

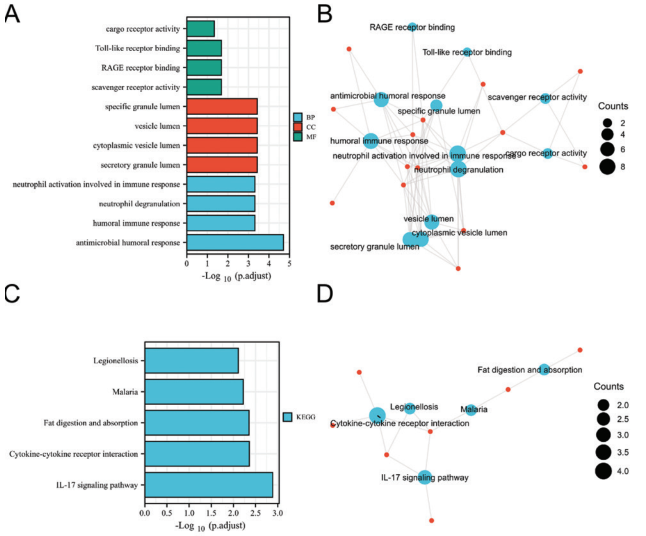


**Figure 5.** (A) 6 Genes from the Core Modules of MCODE. (B) The Top 10 Genes in Critical Places were Eliminated Using the CytoHubba Plug-in in Accordance with MCC Algorithm Guidelines. (C) Venn Diagram of Genes from Groups 1 and 2, where Group 1 Represents a Gene's MCODE Key Module and Group 2 Represents the Top 10 Genes' Important Positions. (D) The PPI Network's Construction and the Various Core Gene Colors.

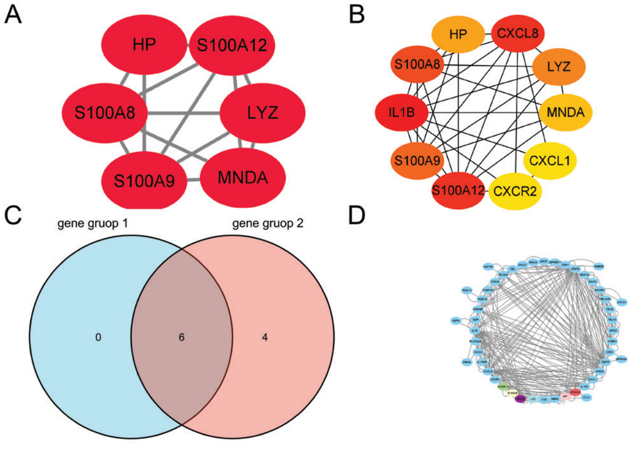
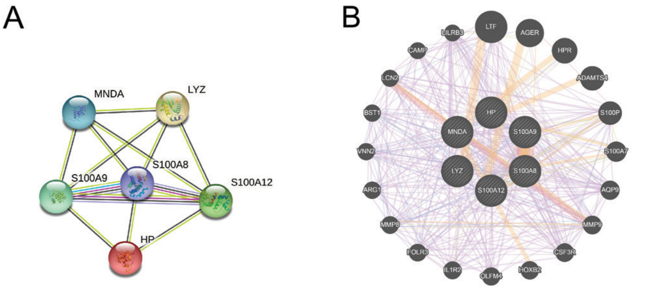


**Table 1.** The Activities of Six Hub Genes

| Gene symbol | Description | Function |
|---|---|---|
| HP | Haptoglobin | Antioxidation |
| MNDA | Myeloid Cell Nuclear Differentiation Antigen | Role in mediating cell death |
| LYZ | Lysozyme | Immune regulation |
| S100A9 | S100 Calcium Binding Protein A9 | Promotes cell proliferation, inhibits cell differentiation and apoptosis. |
| S100A8 | S100 Calcium Binding Protein A8 | Promotes cancer development and tumor metastasis |
| S100A12 | S100 Calcium Binding Protein A12 | Regulates cell function |

**Figure 6.** (A) STRING's Creation of the Top 6 Hub Genes' PPI Network; (B) Analysis of the 6 Hub Genes' PPI Network & Function; The Size of the Circles Reflects the Association with the Input Genes, while the Inner Circles Stand for the Input Genes and the Outside Circles for the Hub Genes Proposed by Gene MANIA.



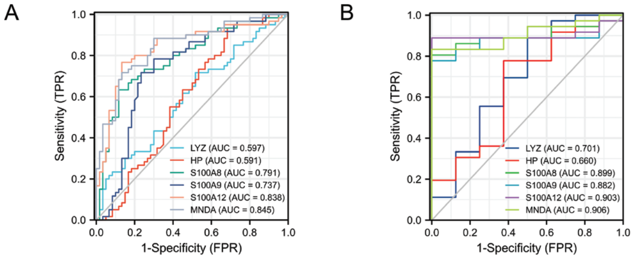
## PPI network analysis, MCODE cluster modules, selection and analysis of hub genes

After the PPI network map was obtained using STRING, a PPI network beautification map was made to separate the hub genes from the common genes, loaded into Cytoscape (version 3.8.2), and exported from Cytoscape. Table 1 displays these hub genes along with further information. These genes make up the majority of the PPI network and may have a substantial role in the etiology of lung cancer and AF. The gene cluster modules were identified using the MCODE plugin. The most significant modules in this network were determined, and the results are displayed in (Figure 5A). Cluster 1 had the greatest cluster rating (5.2, 6 nodes, and 13 edges). Finally, 6 core candidate genes were produced by intersecting the 10 core genes predicted by MCC with the 6 critical module genes acquired by MCODE analysis, namely MNDA, HP, LYZ, S100A9, S100A8, S100A12, (Figure 5C), by using the Venn diagram analysis. GeneMANIA then acquired the PPI network map, which revealed that a total of 6 hub genes had complex PPI networks with 47.30% co-expression, 1.71% co-localization, 18.07% predicted interactions, 4.00% shared protein structural domains, and 28.28% physical interactions. The main functions of the genes included regulation of inflammatory processes, immune response, antioxidant effects, and mediating cell death (Figure 6.).

## Confirmation of hub genes' diagnostic utility

ROC curves were built and AUC was determined related to the levels of these genes' expression in datasets related to lung cancer or atrial fibrillation, S100A8, S100A9, S100A12, as well as MNDA in patients with lung cancer, in order to validate the diagnostic value of the first six hub genes discovered from the aforementioned analysis using normal controls. The diagnostic values obtained were 0.597 [95% CI, 0.495-0.699; $P < .05$], 0.591 (95% CI, 0.487-0.695; $P < .05$), 0.791 (95% CI, 0.709-0.873; $P < .05$), 0.737 (95% CI, 0.643-0.831; $P < .05$), 0.838 (95% CI, 0.762-0.913; $P < .05$), and 0.845 (95% CI, 0.775-0.915; $P < .05$). Figure 12 shows the ROC curve in patients of sinus rhythm and atrial fibrillation. The AUC for 6 hub genes in atrial fibrillation were 0.701

**Figure 7.** The Diagnostic Value of Six Important Genes for Lung Cancer and AF was Examined Using ROC Curve. (A) Analysis of the Lung Cancer ROC Curve. (B) ROC Analysis of the AF Curve.



[95% CI, 0.456-0.946; $P < .05$], 0.660 (95% CI, 0.423-0.897; $P < .05$), 0.899 (95% CI, 0.808-0.990; $P < .05$), 0.882 (95% CI, 0.781-0.983; $P < .05$), 0.903 (95% CI, 0.809-0.997; $P < .05$), and 0.906 (95% CI, 0.819-0.994; $P < .05$) (Figure 7).

## DISCUSSION

Lung cancer, one of the most frequent diseases in the world,[20, 21] and contributes to more than 20% of all cancer-related deaths. Age, obstructive sleep apnea, BMI, hypertension, diabetes mellitus, heart failure, height, myocardial infarction, smoking, and genetic susceptibility are a few established risk factors for the development and duration of AF.[22] In recent years, tumors have also been recognized as a distinct risk factor for the development of AF, notably in lung cancer patients.[23] According to one study, AF is the arrhythmia condition individuals with lung cancer should be most concerned about. Another investigation found that 67.2% of non-small cell lung cancer (NSCLC) patients had cardiovascular comorbidities, of which 7.9% were AF. In this study, approximately 3.8% of NSCLC patients receiving best supportive care had AF, constituting about 7.2% in the treatment group (23/318) and up to 7.6% in the palliative care patient group (22/291).[24] In a survey from the Universal Health Insurance Database in Taiwan, China, 2.3% of 52,089 lung cancer patients had AF compared to 1.3% of non-lung cancer patients, a meaningfully subordinate rate than lung cancer population ($P <0.01$).[25] In addition, AF is a common early postoperative complication of various procedures including lung cancer.[26] Onaitis[10] found in a study of 13906 patients undergoing lung cancer surgery that about 12.6% of them developed postoperative AF.

The lung carcinoma and AF databases in GEO were searched in this study, and an aggregate of 49 DEGs were found among both disorders. To investigate the relationships among the DEGs, we also carried out GO enrichment and KEGG pathway enrichment studies. These pathways are also implicated in neurobiology, cancer, and many other illnesses.[27] Through the recruitment of the aforementioned molecules as well as activation of signaling pathways such as the RAGE, a wide range of intracellular signaling molecules get activated, including transcription factors NF-κB, MAP kinases, as well as adhesion molecules, among a number of others, eventually encouraging tumorigenesis.[28, 29] The mechanism of lung carcinogenesis is also closely related to the development of a persistent inflammatory response. In addition, an additional pathophysiological mechanism underlying the onset of AF has been proposed: the overexpression of receptors for advanced glycosylation endproducts (RAGE). Pro-fibrotic and pro-inflammatory pathways are known to enhance vulnerability to AF, however the precise molecular processes behind atrial remodeling are not well understood.[30] By interacting with the cellular advanced glycosylation endproducts (AGE) receptor RAGE, which promotes fibrosis and inflammation by sustained activation of nuclear factor kappa-B, upregulation of cytokines and adhesion molecules, and AGEs, which are non-enzymatic glycosylation as well as oxidation byproducts of proteins as well as lipids, may lead to the onset of AF and contribute to its pathophysiological progression.[31, 32] Thus, RAGE contributes to the co-morbidity of lung cancer and AF through the common pathway of the collective induction of inflammatory responses.

Additionally, tumor cells produce toll-like receptors (TLRs), a type of single-channel transmembrane proteins, which, mostly through the inflammatory response, play a crucial function in the innate immune system. Humans possess ten TLR species (TLR1–TLR10), and it is important to note that TLR2, TLR4, TLR5, TLR7–8, and TLR9 are more abundantly expressed in lung cancer tissues than in healthy lung tissues.[33, 34] TLRs contribute significantly to the immune response against tumor cells by preventing the immune system from destroying a crucial cancer growth marker. The inflammatory response is coordinated by TLR-activated downstream signaling pathways, and it has been discovered that there is a clear connection between chronic inflammation and cancer.[35] An infection or damage triggers inflammation, which is a basic immunological response. By secreting cytokines, chemokines, growth factors, reactive oxygen species (ROS), and reactive nitrogen species (RNS), innate immune cells such macrophages, mast cells, dendritic cells, and natural killer cells can heal tissue damage and eradicate invading pathogenic microorganisms.[36] The aforementioned inflammatory mediators also cause genetic and epigenetic changes, such as point mutations in tumor suppressor genes, DNA methylation, and post-translational modifications, which alter some crucial pathways necessary to maintain normal cellular endocytosis and ultimately promote the growth and spread of cancer.[37, 38] Inflammation has a key role in the etiology of AF, as is now known. TLRs are pattern recognition receptors that alert the body's innate immune system to the presence of pathogens by detecting certain molecular patterns. According to one research, patients with non-valvular AF have more TLR-2 expression than volunteers with sinus rhythm.[40] An increase in TLR-2 gene expression and protein content was observed in the right heart ear specimens from patients with persistent as well as paroxysmal AF following valve replacement alone. The authors hypothesized that the activated TLR pathway may recruit downstream IRAK4/1 through a Mal/myd88-dependent mechanism, ultimately leading to the activation of activated

B-cell kappa-light chain-enhancing factors. The associated production of pro-inflammatory cytokines and the pro-inflammatory state leads to structural and electrophysiological changes in the heart that are closely associated with the development and maintenance of AF.[41]

According to the KEGG pathway enrichment analysis, DEGs were significantly enriched in immune-related pathways and cytokine-cytokine receptor interactions. Six structurally similar cytokines, spanning from IL-17A to IL-17F, make up the interleukin 17 (IL-17) cytokine family. This pleiotropic cytokine is crucial in many chronic inflammatory disorders.[42] The progenitors of granulocytes, macrophages, dendritic cells, and various other immune cells are called myeloid-derived suppressor cells (MDSC). The recruitment of MDSC is facilitated by IL-17 production in the tumor microenvironment, which inhibits the body's antitumor immune response.[43] IL-17 contributes to the growth of lung cancer by directly or indirectly promoting tumor angiogenesis, cell proliferation, and apoptosis while inhibiting anti-tumor immune responses via activating inflammatory signaling pathways. Furthermore, it was shown that IL-17, by activating the epithelial mesenchymal transition (EMT) in lung cancer cells (A549), had a significant role in the genesis and spread of the disease (Ki-67 test). These effects were found to promote the proliferation of A549 as well as SK-MES-1 cells as well as inhibit apoptosis. Atrial fibrillation is also significantly influenced by IL-17. Atrial activation occurs rapidly and irregularly in AF, which lowers the cardiac output. The elevated atrial leukocyte infiltration reported in both solitary and structural AF is thought to be predictive of the postoperative AF. IL-17A can directly promote the release of cytokines that cause inflammation (IL-1β, IL-6, etc.), thus increasing atrial fibrosis and altered electrical signals and promoting the development of AF.[45, 46]

In addition, we screened via PPI network, 10 candidate genes, LYZ, HP, S100A8, S100A9, S100A12, MNDA, IL1B, CXCL8, CXCL1, CXCR2, and further. By using the MCODE plug-in to find gene clustering modules, we were able to group the 10 candidate genes identified by PPI into four modules with the best score. These four modules were then intersected to produce the six core genes LYZ, HP, S100A8, S100A9, S100A12, and MNDA. To verify the diagnostic usefulness to individuals with lung cancer and AF, we chose 6 hub genes ($P$ <0.05). From the above table 1, it can be obtained that in predicting lung cancer and lung normal outcomes, the variables LYZ and HP had low accuracy in predicting, while the variables S100A8, S100A9, S100A12, and MNDA had predictive ability with varying accuracy, among which S100A12 and MNDA had high accuracy in predicted power. These genes may have a significant role in predicting the likelihood of developing lung cancer and AF.

Myeloid cells like monocytes and neutrophils release the calgranulin proteins S100A8 (calgranulin A), S100A9 (calgranulin B), and S100A12 (calgranulin C). They are frequently co-expressed as S100A8/A9 heterodimer dimer (calprotectin), and have been described as important regulators of immune response and inflammation.[47, 48] S100A8/A9 has been linked to the growth and progression of only a few types of tumors such as breast cancer, lung cancer, and colon tumors.[49] Further research revealed that S100A8/A9 in lung cancer binds to TLR4 and RAGE on macrophages and activates the NF-κB signaling pathway, which causes the production of pro-inflammatory cytokines. Subsequently, the pro-inflammatory cytokines promote the recruitment of more monocytes and neutrophils, producing a positive feedback loop. S100A8 and S100A9 are also chemo-attractants for myeloid-derived suppressor cells (MDSC). By reducing the cytotoxicity of cytotoxic CD8+ T cells and NK cells, MDSC that have been transported from bone marrow to peripheral blood dampen the immune system and promote carcinogenesis.[50, 51] Additionally, bone marrow adipocytes (BMAs) are involved in non-anchored growth, motility, and invasion by activating S100A8/A9, while S100A12 binds to RAGE followed by signaling through NF-κB as well as MAPK pathways. Neurofibril (NPTN) is also involved in non-anchored growth, motility, and invasion by binding S100A8/A9 and then mediating the downstream signaling pathway TRAF2/RAS-NFIA, enhancing intercellular adhesion molecule-1 (ICAM-1), vascular cell adhesion molecule-1 (VCAM-1), nuclear factor κB and tumor necrosis factor (TNF-α ) expression, which directly causes the activation of mononuclear phagocytes, lymphocytes, and endothelial cells, and as a result leads to the synthesis and secretion of pro-inflammatory cytokines.[52] Cardiovascular illnesses including atherosclerosis, heart failure, and acute coronary syndrome are all intimately connected with S100A8/A9 and S100A12, although research on their link to atrial fibrillation is lacking. By promoting the production of pro-inflammatory mediators and enhancing the inflammatory response, the Toll-like receptor 4 (TLR4) and the receptor for advanced end products of glycation (RAGE), to which S100A8/A9 binds, play a role in the initiation and development of immune inflammatory disorders. This association between S100A8/A9 and AF is due to this. The aforementioned processes also contribute to the development of atrial fibrillation. We thus postulate that increased levels of S100A8/A9 and S100A12 are connected to the onset of AF, although more experimental confirmation is required for the limited research that have been conducted in this area.

In the serum globin-2 fraction, haptoglobin (Hp) is an acidic glycoprotein that comes in three basic genotypes: Hp1-1, Hp2-1, and Hp2-2. Each genotype has a unique biological purpose. Its main role is to bind hemoglobin and prevent oxidative stress, immune response, and promote angiogenesis. Its levels have been found to be significantly elevated in plasma in infections, inflammatory conditions, and various malignancies, includes leukemia, breast cancer, esophageal squamous cell carcinoma, and lung and bladder cancer[53] It has been found that Hp and its variants exert inhibitory effects on T-lymphocyte function, mainly through T-cell and inhibit T-helper cells by down-regulating cytokine production and related mechanisms, and this inhibition helps to protect tumors from immune attack, enabling tumor cell evasion from surveillance mechanisms.[54]

In summary, we can find that Hp is involved in the process of tumorigenesis through the immune response pathway, and there are few studies on the role of Hp in AF. Nonetheless, we can speculate that the mechanism of Hp linked to immune response may also be involved in the development of AF. Lung cancer and AF have significant levels of LYZ and MNAD expression; however, additional research is needed to determine the correlation between LYZ and MNAD in lung cancer and AF.

## CONCLUSIONS

In conclusion, we effectively obtained an insight into the molecular changes involved in the pathogenesis of lung tumors as well as AF. We have also identified several potential diagnostic and therapeutic targets, including LYZ, HP, S100A8, S100A9, S100A12, and MDAD using PPI analysis of differential gene expression analysis, functional enrichment analysis, and PPI analysis of DEGs in lung cancer and AF. These core genes are mainly enriched in immune related biological pathways and functions. The results of this study may provide new understanding of the genetic pathogenesis of lung cancer and AF.

## ABBREVIATIONS

AF, Atrial Fibrillation; DEGs: Differentially expressed genes; GEO, Gene Expression Omnibus; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, Protein–protein interaction; co-DEGs, Co-expressed differentially expressed genes; CC, Cellular component; MF, Molecular function; BP, Biological process; ROC, Receiver operator characteristic; AUC, Area under the curve; RAGE, Receptor for advanced glycation end products; TLRs, toll-like receptors.

## ETHICS APPROVAL

According to the Helsinki Declaration, the study has received approval from the affiliated hospital of North Sichuan Medical College ethics committee. Every piece of information came from a public database.

## DATA AVAILABILITY

The [Gene Expression Omnibus databased] repository/database has the datasets used in the current investigation available for public use. [https://www.ncbi.nlm.nih.gov/geo/], the Search Tool for the Retrieval of Interacting Genes[http://string-db.org], the [GeneMANIA] online database, [ https://genemania.org/], and Xian Tao, [https://www.xiantao.love/].

## AUTHOR DISCLOSURE STATEMENT

It is stated by the authors that they have no competing interests. The final manuscript was reviewed and approved by all writers.

## ACKNOWLEDGEMENT

We are appreciative to the participants for supplying the data for our study and setting up data collecting.

## AUTHOR CONTRIBUTIONS

As co-first authors Dongmei Yang and Yong Chen have contributed equally to this study. This study was conceived of and designed by Yong Chen and Dongmei Yang. The idea, formal analysis, and first draught writing were completed by Yonglin Yu and Xiaoju Chen. Yong Chen and Yonglin Yu oversaw the investigation. The paper was written by Yong Chen and Dongmei Yang. Resources and data curation were handled using Dongmei Yang.

## FUNDING

This study was supported by the 2020 Municipal Applied Technology Research and Development Fund Project (no. 20YFZJ0105).

## REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. [J]. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
2. Roth GA, Mensah GA, Johnson CO, et al; GBD-NHLBI-JACC Global Burden of Cardiovascular Diseases Writing Group. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. [J]. *J Am Coll Cardiol*. 2020;76(25):2982-3021. doi:10.1016/j.jacc.2020.11.010
3. Zubair Khan M, Gupta A, Patel K, et al. Association of atrial fibrillation and various cancer subtypes. [J]. *J Arrhythm*. 2021;37(5):1205-1214. doi:10.1002/joa3.12589
4. Ball J, Carrington MJ, McMurray JJ, Stewart S. Atrial fibrillation: profile and burden of an evolving epidemic in the 21st century. [J]. *Int J Cardiol*. 2013;167(5):1807-1824. doi:10.1016/j.ijcard.2012.12.093
5. Jakobsen CB, Lamberts M, Carlson N, et al. Incidence of atrial fibrillation in different major cancer subtypes: a Nationwide population-based 12 year follow up study. [J]. *BMC Cancer*. 2019;19(1):1105. doi:10.1186/s12885-019-6314-9
6. Yun JP, Choi EK, Han KD, et al. Risk of Atrial Fibrillation According to Cancer Type: A Nationwide Population-Based Study. [J]. *JACC Cardiooncol*. 2021;3(2):221-232. doi:10.1016/j.jaccao.2021.03.006
7. Yang PS, Shim M, Kang SH, et al. Incidence of atrial fibrillation in cancer patients according to cancer type [J]. *European Heart Journal*. 2020;41(Supplement_2):ehaa946.3271 . doi:10.1093/ehjci/ehaa946.3271
8. Hu YF, Liu CJ, Chang PM, et al. Incident thromboembolism and heart failure associated with new-onset atrial fibrillation in cancer patients. [J]. *Int J Cardiol*. 2013;165(2):355-357. doi:10.1016/j.ijcard.2012.08.036
9. Chu G, Versteeg HH, Verschoor AJ, et al. Atrial fibrillation and cancer - An unexplored field in cardiovascular oncology. [J]. *Blood Rev*. 2019;35:59-67. doi:10.1016/j.blre.2019.03.005
10. Onaitis M, D'Amico T, Zhao Y, O'Brien S, Harpole D. Risk factors for atrial fibrillation after lung cancer surgery: analysis of the Society of Thoracic Surgeons general thoracic surgery database. [J]. *Ann Thorac Surg*. 2010;90(2):368-374. doi:10.1016/j.athoracsur.2010.03.100
11. Albini A, Pennesi G, Donatelli F, Cammarota R, De Flora S, Noonan DM. Cardiotoxicity of anticancer drugs: the need for cardio-oncology and cardio-oncological prevention. [J]. *J Natl Cancer Inst*. 2010;102(1):14-25. doi:10.1093/jnci/djp440
12. Hu YF, Chen YJ, Lin YJ, Chen SA. Inflammation and the pathogenesis of atrial fibrillation. [J]. *Nat Rev Cardiol*. 2015;12(4):230-243. doi:10.1038/nrcardio.2015.2
13. Farmakis D, Parissis J, Filippatos G. Insights into onco-cardiology: atrial fibrillation in cancer. [J]. *J Am Coll Cardiol*. 2014;63(10):945-953. doi:10.1016/j.jacc.2013.11.026
14. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. [J]. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007
15. Gene Ontology Consortium. The Gene Ontology (GO) project in 2006. [J]. *Nucleic Acids Res*. 2006;34(Database issue):D322-D326.
16. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. [J]. *Nucleic Acids Res*. 2016;44(D1):D457-D462. doi:10.1093/nar/gkv1070
17. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. [J]. *Nucleic Acids Res*. 2021;49(D1):D605-D612. doi:10.1093/nar/gkaa1074
18. Chami M, Oulès B, Paterlini-Bréchot P. Cytobiological consequences of calcium-signaling alterations induced by human viral proteins. [J]. *Biochim Biophys Acta*. 2006;1763(11):1344-1362. doi:10.1016/j.bbamcr.2006.09.025
19. Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function [J]. *Nucleic Acids Res*. 2010;38(Web Server issue):W214-220. doi:10.1093/nar/gkq537
20. O'Dowd E, Mackenzie J, Balata H. Lung cancer for the non-respiratory physician. [J]. *Clin Med (Lond)*. 2021;21(6):e578-e583. doi:10.7861/clinmed.2021-0657
21. Benjamin EJ, Muntner P, Alonso A, et al; American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2019 Update: A Report From the American Heart Association. [J]. *Circulation*. 2019;139(10):e56-e528. doi:10.1161/CIR.0000000000000659
22. Staerk L, Sherer JA, Ko D, Benjamin EJ, Helm RH. Atrial Fibrillation: Epidemiology, Pathophysiology, and Clinical Outcomes. [J]. *Circ Res*. 2017;120(9):1501-1517. doi:10.1161/CIRCRESAHA.117.309732
23. Boriani G, Lee G, Parrini I, et al; Council of Cardio-Oncology of the European Society of Cardiology. Anticoagulation in patients with atrial fibrillation and active cancer: an international survey on patient management. [J]. *Eur J Prev Cardiol*. 2021;28(6):611-621. doi:10.1093/eurjpc/zwaa054
24. Yuan M, Li QG. Lung Cancer and Risk of Cardiovascular Disease: A Meta-analysis of Cohort Studies. [J]. *J Cardiothorac Vasc Anesth*. 2018;32(1):e25-e27. doi:10.1053/j.jvca.2017.04.033
25. Jegou D, Dubois C, Schillemans V, et al. Use of health insurance data to identify and quantify the prevalence of main comorbidities in lung cancer patients. [J]. *Lung Cancer*. 2018;125:238-244. doi:10.1016/j.lungcan.2018.10.002
26. Semeraro GC, Meroni CA, Cipolla CM, Cardinale DM. Atrial Fibrillation after Lung Cancer Surgery: Prediction, Prevention and Anticoagulation Management. [J]. *Cancers (Basel)*. 2021;13(16):4012. doi:10.3390/cancers13164012
27. Mukherjee TK, Malik P, Hoidal JR. Receptor for Advanced Glycation End Products (RAGE) and Its Polymorphic Variants as Predictive Diagnostic and Prognostic Markers of NSCLCs: a Perspective. [J]. *Curr Oncol Rep*. 2021;23(1):12. doi:10.1007/s11912-020-00992-x
28. Rojas A, Schneider I, Lindner C, Gonzalez I, Morales MA. The RAGE/multiligand axis: a new actor in tumor biology. [J]. *Biosci Rep*. 2022;42(7):BSR20220395. doi:10.1042/BSR20220395
29. Rojas A, Delgado-López F, González I, Pérez-Castro R, Romero J, Rojas I. The receptor for advanced glycation end-products: a complex signaling scenario for a promiscuous receptor. [J]. *Cell Signal*. 2013;25(3):609-614. doi:10.1016/j.cellsig.2012.11.022
30. Ihara K, Sasano T. Role of Inflammation in the Pathogenesis of Atrial Fibrillation. [J]. *Front Physiol*. 2022;13:862164. doi:10.3389/fphys.2022.862164
31. Ramasamy R, Vannucci SJ, Yan SS, Herold K, Yan SF, Schmidt AM. Advanced glycation end products and RAGE: a common thread in aging, diabetes, neurodegeneration, and inflammation. [J]. *Glycobiology*. 2005;15(7):16R-28R. doi:10.1093/glycob/cwi053
32. Lancefield TF, Patel SK, Freeman M, et al. The Receptor for Advanced Glycation End Products (RAGE) Is Associated with Persistent Atrial Fibrillation. [J]. *PLoS One*. 2016;11(9):e0161715. doi:10.1371/journal.pone.0161715
33. Arora S, Ahmad S, Irshad R, et al. TLRs in pulmonary diseases. [J]. *Life Sci*. 2019;233:116671. doi:10.1016/j.lfs.2019.116671
34. Gu J, Liu Y, Xie B, Ye P, Huang J, Lu Z. Roles of toll-like receptors: from inflammation to lung cancer progression. [J]. *Biomed Rep*. 2018;8(2):126-132. doi:10.3892/br.2017.1034
35. Broz P, Monack DM. Newly described pattern recognition receptors team up against intracellular pathogens. [J]. *Nat Rev Immunol*. 2013;13(8):551-565. doi:10.1038/nri3479
36. Fernandes JV, Cobucci RN, Jatobá CA, Fernandes TA, de Azevedo JW, de Araújo JM. The role of the mediators of inflammation in cancer development. [J]. *Pathol Oncol Res*. 2015;21(3):527-534. doi:10.1007/s12253-015-9913-z
37. Yasmin R, Siraj S, Hassan A, Khan AR, Abbasi R, Ahmad N. Epigenetic regulation of inflammatory cytokines and associated genes in human malignancies. [J]. *Mediators Inflamm*. 2015;2015:201703. doi:10.1155/2015/201703
38. Florea ID, Karaoulani C. Epigenetic Changes of the Immune System with Role in Tumor Development. [J]. *Methods Mol Biol*. 2018;1856:203-218. doi:10.1007/978-1-4939-8751-1_11
39. Gurses KM, Kocyigit D, Yalcin MU, et al. Monocyte Toll-Like Receptor Expression in Patients With Atrial Fibrillation. [J]. *Am J Cardiol*. 2016;117(9):1463-1467. doi:10.1016/j.amjcard.2016.02.014

40. Ichiki H, Orihara K, Hamasaki S, et al. The role of infection in the development of non-valvular atrial fibrillation: up-regulation of Toll-like receptor 2 expression levels on monocytes. [J]. *J Cardiol*. 2009;53(1):127-135. doi:10.1016/j.jjcc.2008.09.008
41. Wang J, Xue L, Cao H, Cui F, Dai T, Chen Y. TLR2 was overexpressed independent of IL-6 in patients with valvular atrial fibrillation. [J]. *J Biomed Res*. 2011;25(3):178-184. doi:10.1016/S1674-8301(11)60023-7
42. Chung SH, Ye XQ, Iwakura Y. Interleukin-17 family members in health and disease. [J]. *Int Immunol*. 2021;33(12):723-729. doi:10.1093/intimm/dxab075
43. Chang SH, Mirabolfathinejad SG, Katta H, et al. T helper 17 cells play a critical pathogenic role in lung cancer. [J]. *Proc Natl Acad Sci USA*. 2014;111(15):5664-5669. doi:10.1073/pnas.1319051111
44. Gu K, Li MM, Shen J, et al. Interleukin-17-induced EMT promotes lung cancer cell migration and invasion via NF-κB/ZEB1 signal pathway. [J]. *Am J Cancer Res*. 2015;5(3):1169-1179.
45. Liuzzo G, Trotta F, Pedicino D. Interleukin-17 in atherosclerosis and cardiovascular disease: the good, the bad, and the unknown [J]. *Eur Heart J*. 2013;34(8):556-559. oi: doi:10.1093/eurheartj/ehs399
46. Valente AJ, Yoshida T, Gardner JD, Somanna N, Delafontaine P, Chandrasekar B. Interleukin-17A stimulates cardiac fibroblast proliferation and migration via negative regulation of the dual-specificity phosphatase MKP-1/DUSP-1. [J]. *Cell Signal*. 2012;24(2):560-568. doi:10.1016/j.cellsig.2011.10.010
47. Gebhardt C, Németh J, Angel P, Hess J. S100A8 and S100A9 in inflammation and cancer. [J]. *Biochem Pharmacol*. 2006;72(11):1622-1631. doi:10.1016/j.bcp.2006.05.017
48. Ehrchen JM, Sunderkötter C, Foell D, Vogl T, Roth J. The endogenous Toll-like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection, autoimmunity, and cancer. [J]. *J Leukoc Biol*. 2009;86(3):557-566. doi:10.1189/jlb.1008647
49. Wang T, Du G, Wang D. The S100 protein family in lung cancer. [J]. *Clin Chim Acta*. 2021;520:67-70. doi:10.1016/j.cca.2021.05.028
50. Lim MY, Thomas PS. Biomarkers in Exhaled Breath Condensate and Serum of Chronic Obstructive Pulmonary Disease and Non-Small-Cell Lung Cancer. [J]. *Int J Chronic Dis*. 2013;2013:578613. doi:10.1155/2013/578613
51. Wang S, Song R, Wang Z, Jing Z, Wang S, Ma J. S100A8/A9 in Inflammation. [J]. *Front Immunol*. 2018;9:1298. doi:10.3389/fimmu.2018.01298
52. Yan WX, Armishaw C, Goyette J, et al. Mast cell and monocyte recruitment by S100A12 and its hinge domain. [J]. *J Biol Chem*. 2008;283(19):13035-13043. doi:10.1074/jbc.M710388200
53. Abdullah M, Schultz H, Kähler D, et al. Expression of the acute phase protein haptoglobin in human lung cancer and tumor-free lung tissues. [J]. *Pathol Res Pract*. 2009;205(9):639-647. doi:10.1016/j.prp.2009.04.007
54. Arredouani M, Matthijs P, Van Hoeyveld E, et al. Haptoglobin directly affects T cells and suppresses T helper cell type 2 cytokine release. [J]. *Immunology*. 2003;108(2):144-151. doi:10.1046/j.1365-2567.2003.01569.x