

Research Article

Comprehensive Analysis of the Molecular Characteristics and Prognosis value of AT II-associated Genes in Non-small Cell Lung Cancer

Liping Ren ¹, Xiaoxia Wen ², Mujiexin Liu ³, Yao Xiao ⁴, Ping Leng ²,
Huaichao Luo ⁴, Pei Tao ⁵, and Lei Xie ⁶

¹School of Healthcare Technology, Chengdu Neusoft University, Chengdu, China

²Chongqing Key Laboratory of Sichuan-Chongqing Co-construction for Diagnosis and Treatment of Infectious Diseases Integrated Traditional Chinese and Western Medicine, College of Medical Technology, Chengdu University of Traditional Chinese Medicine, Chengdu, China

³Ineye hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, China

⁴Department of clinical laboratory, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu, Sichuan, China

⁵Chengdu Women's and Children's Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Sichuan 611731, China

⁶The Sichuan Provincial Key Laboratory for Human Disease Gene Study and Department of Laboratory Medicine, Sichuan Provincial People's Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China

Correspondence should be addressed to Liping Ren; 273982408@qq.com, Pei Tao; 9944919@qq.com, and Lei Xie; xielei@uestc.edu.cn

Received 7 June 2022; Accepted 10 September 2022; Published 26 September 2022

Academic Editor: Leyi Wei

Copyright © 2022 Liping Ren et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Alveolar type II (AT II) is a key structure of the distal lung epithelium and essential to maintain normal lung homeostasis. Dedifferentiation of AT II cells is significantly correlated with lung tumor progression. However, the potential molecular mechanism and clinical significance of AT II-associated genes for lung cancer has not yet been fully elucidated. In this study, we comprehensively analyzed the gene expression, prognosis value, genetic alteration, and immune cell infiltration of eight AT II-associated genes (AQP4, SFTP4, SFTP5, SFTP6, CLDN18, FOXA2, NKX2-1, and PGC) in Non-small Cell Lung Cancer (NSCLC). The results have shown that the expression of eight genes were remarkably reduced in cancer tissues and observably relating to clinical cancer stages. Survival analysis of the eight genes revealed that low-expression of CLDN18, FOXA2, NKX2-1, PGC, SFTP4, SFTP5, and SFTP6 were significantly related to a reduced progression-free survival (FP), and low CLDN18, FOXA2, and SFTP6 mRNA expression led to a short postprogression survival (PPS). Meanwhile, the alteration of 8 AT II-associated genes covered 273 out of 1053 NSCLC samples (26%). Additionally, the expression level of eight genes were significantly correlated with the infiltration of diverse immune cells, including six types of CD4+T cells, macrophages, neutrophils, B cells, CD8+ T cells, and dendritic cells. In summary, this study provided clues of the values of eight AT II-associated genes as clinical biomarkers and therapeutic targets in NSCLC and might provide some new inspirations to assist the design of new immunotherapies.

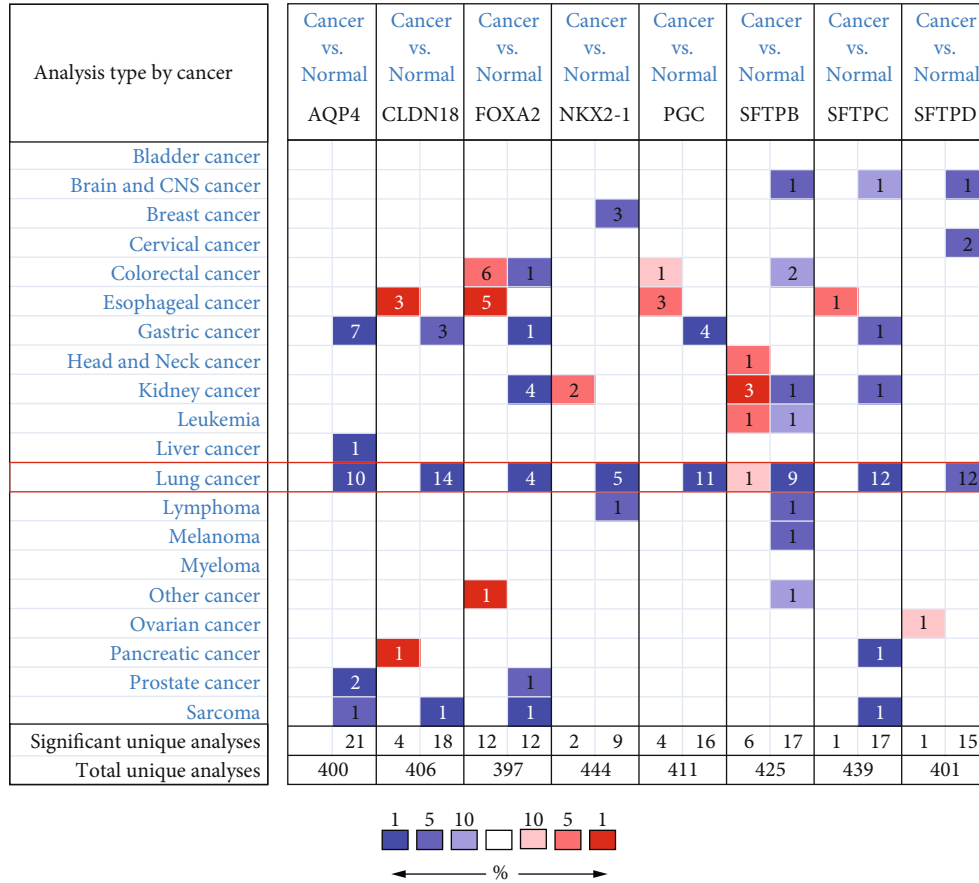


FIGURE 1: Expression of AT II-associated genes in different cancer types (Oncomine). The graphic demonstrates the numbers of datasets with statistically significant alterations in the mRNA expression of the target gene: upregulated (red) and downregulated (blue). The following criteria were used: p value of 0.05, fold change of 2, and gene rank of 10%. As shown in the green frame, expression levels of AQP4, SFTPB, SFTPC, SFTPD, CLDN18, FOXA2, NKX2-1, and PGC were significantly reduced in lung cancer.

1. Introduction

Lung cancer is one of the most commonly diagnosed cancers and the leading cause of cancer-related death in the world [1–3]. Nonsmall cell lung cancer (NSCLC) is one of the most majorly types of lung cancer (approximately 85%), mainly including lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) [2, 4–7]. Studies have demonstrated that the average five-year survival rate of NSCLC patients is 15% [1, 8]. This poor survival rate is attributable to many factors, such as delays in the diagnosis of lung cancer and limited therapies currently available [9, 10]. Over the past decade, with the improvement of treatment technologies and the emergence of the era of precision radiotherapy, the diagnosis, and treatment of lung cancer have been improved to a certain extent [9, 11–16]. Despite advances in treatment, the overall prognosis for NSCLC has not yet improved significantly.

The alveolar cells are mainly composed of alveolar type I (AT I) cells and alveolar type II (AT II) cells [17, 18]. There into, AT II is a key structure of the distal lung epithelium and has a secretory function that is essential to maintaining normal lung homeostasis [19]. In recent years, there is currently substantial evidence showing that AT II and AT II-

associated genes are significantly related to the occurrence and development of multiple diseases [20]. One of the pathological features of the idiopathic pulmonary fibrosis (IPF) lung is the senescence of AT II [21, 22]. AT II is also involved in the occurrence and development of Chronic obstructive pulmonary disease (COPD) through the upregulated expression of many anti- or proinflammatory genes, including genes encoding oxygenase 2 (HO-2) and inducible nitric oxidase (iNOS) [20]. Importantly, several studies have also shown that AT II plays a crucial role in the oncogenesis of lung cancer [8, 23]. Single-cell RNA sequencing of lung cancer tissues revealed that some AT II-associated genes expressed differently in the lung cancer cells, including aquaporin 4 (AQP4), surfactant pulmonary associated protein B (SFTPB), surfactant pulmonary associated protein C (SFTPC), surfactant pulmonary associated protein D (SFTPD), claudin 18 (CLDN18), forkhead box A2 (FOXA2), NKX homeobox-1 gene (NKX2-1), and pepsinogen C (PGC) [24]. However, the potential values of these AT II cell-related genes in NSCLC have not been fully clarified.

Therefore, in this study, we performed a comprehensive analysis, including analysis of gene expression, prognosis value, genetic alteration, and immune cell infiltration of these eight AT II-associated genes in two subtypes of

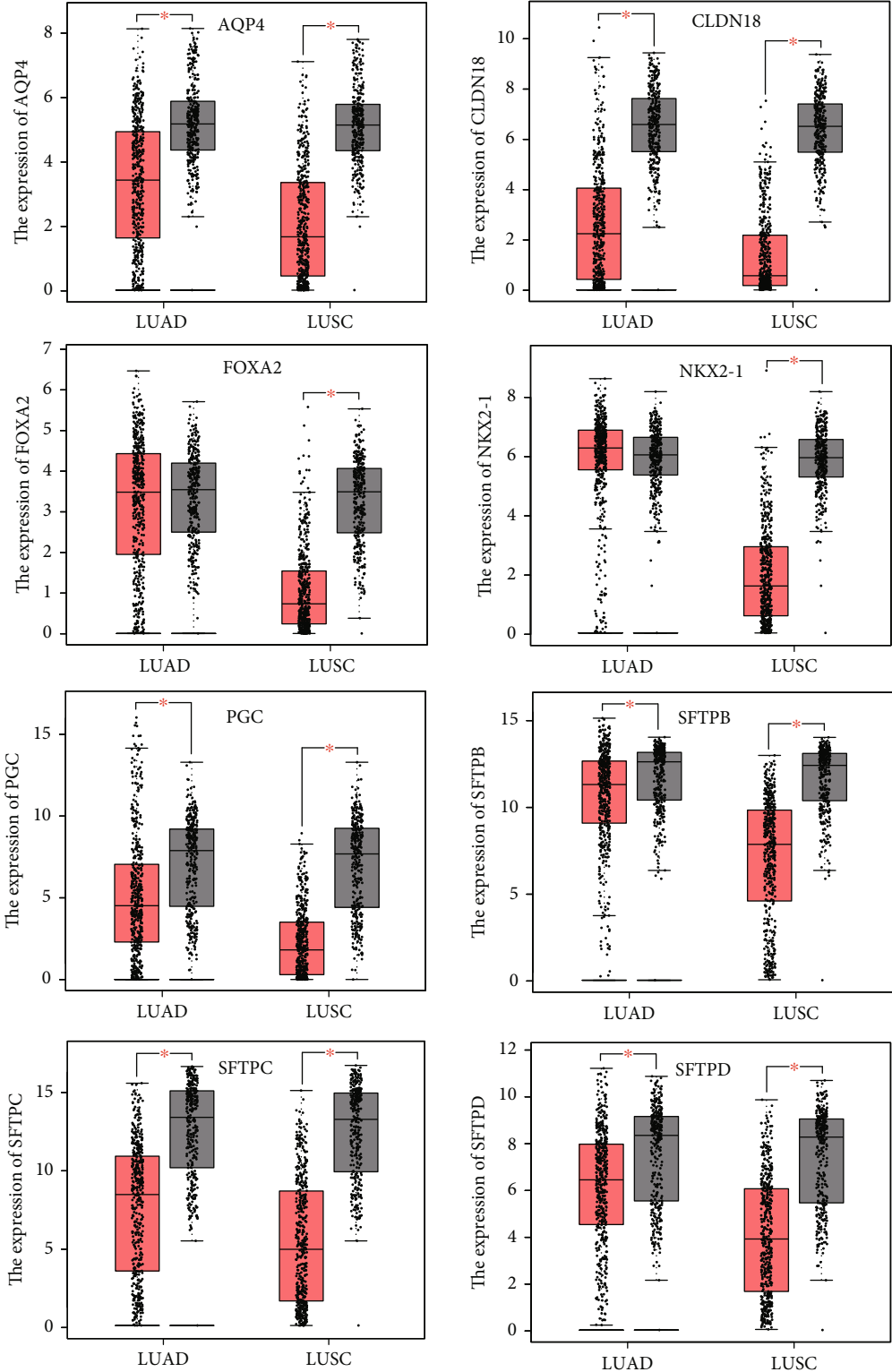


FIGURE 2: The expressions of the AT II-associated genes in LUAD and normal tissues and LUSC and normal tissues (GEPIA). The results indicated that AQP4, CLDN18, PGC, SFTPB, SFTPC, and SFTPD were lower in LUAD tissues than in normal tissue, and AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD were lower in the LUSC tissues than the normal tissues. $*p < 0.01$.

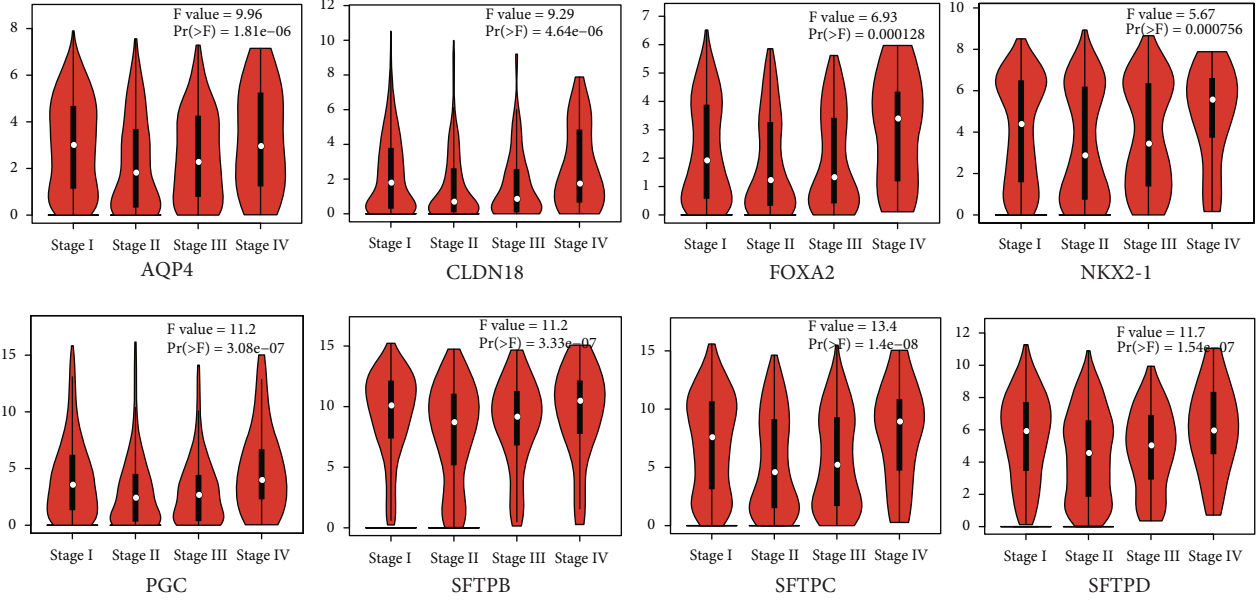


FIGURE 3: Correlation between expression of AT II-associated genes and tumor stage in NSCLC (GEPiA). The expressions of AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD were distinctly related to patients' individual cancer stages (one-way ANOVA).

TABLE 1: The relationship between the expression level of AT II-associated genes and NSCLC prognosis.

AT II-associated genes	Kaplan-Meier plotter (log rank p)			GEPIA (log rank p)
	OS	FP	PPS	DFS
AQP4	2.40E-04	5.60E-02	4.00E-01	7.40E-01
CLDN18	1.90E-05	9.10E-04	3.20E-02	6.90E-01
FOXA2	1.60E-12	6.70E-05	2.10E-02	8.10E-01
NKX2-1	4.90E-10	3.10E-02	1.20E-01	1.30E-01
PGC	1.00E-08	2.40E-04	5.90E-01	2.20E-01
SFTP-B	6.30E-10	4.80E-02	3.80E-01	4.20E-01
SFTPC	1.40E-03	4.00E-02	1.10E-01	2.90E-01
SFTPD	1.60E-10	6.20E-05	2.10E-02	4.00E-01

Note: OS: overall survival; FP: progression-free survival; PPS: postprogression survival; DFS: disease-free survival.

NSCLC (LUAD and LUSC). It aims to provide clinicians with additional information to assess and adjust the diagnostic methods and treatment options of NSCLC patients.

2. Materials and Methods

2.1. Oncomine. Oncomine database is a publicly accessible online cancer microarray database. (<http://www.oncomine.org/>), which provides a genome-wide expression analysis for a wide variety of tumor types [25]. In this study, it was utilized to analyze the transcription levels of AT II-associated genes in NSCLC tissues and their corresponding adjacent normal control samples. The p value < 0.05 (Student's t -test), fold change of 2, and gene rank in the top 10% were set as the significance.

2.2. Gene Expression Profiling Interactive Analysis (GEPiA). GEPiA (<http://gepia.cancer-pku.cn/index.html>) is a newly developed interactive web server for analyzing the RNA sequencing expression data of 9736 tumors and 8587 normal

samples from the TCGA and Genotype-Tissue Expression dataset [26]. GEPiA offers customizable functions such as tumor/normal differential expression analysis, patient survival analysis, similar gene detection, correlation analysis, and dimensionality reduction analysis. The Student's t -test was used to generate a p value (p value < 0.01). In this study, we performed the pathological type and stage analysis of eight AT II-associated genes using the "LUAD" and "LUSC" datasets (one-way ANOVA).

2.3. Kaplan-Meier Plotter. Kaplan-Meier Plotter (<https://kmplot.com/analysis/>) is a useful prognostic biomarker assessment tool that can assess the effect of 54k genes on survival in 21 cancer types [27]. In this study, LUAD and LUSC patients were split into high and low-expression groups based on median values of AT II-associated genes expression and analyze their prognostic value in LUAD and LUSC regarding OS (overall survival), FP (first progression), and PPS (postprogression survival). The hazard ratio

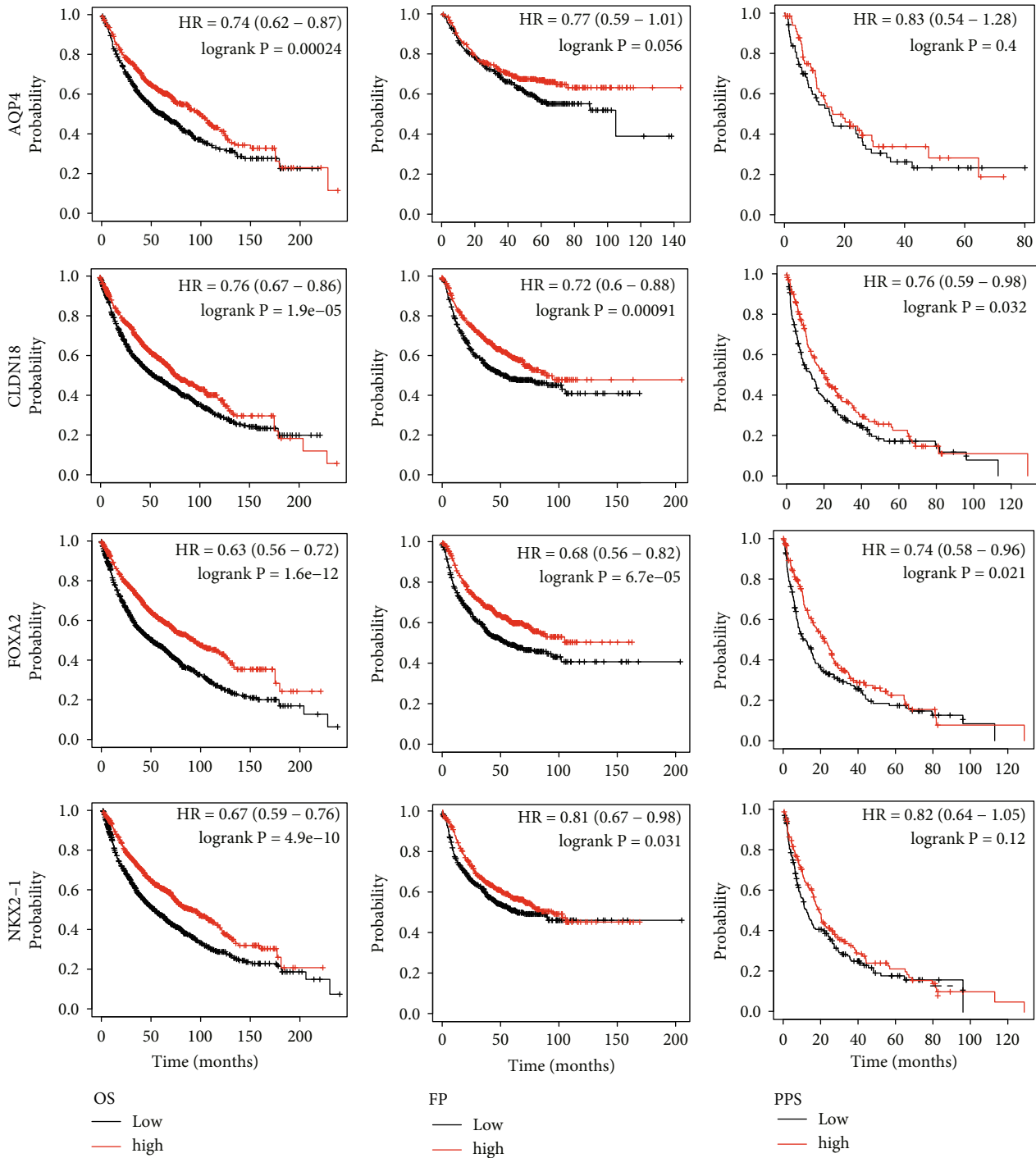


FIGURE 4: Prognostic value of AT II-associated genes (AQP4, CLDN18, FOXA2, and NKX2-1) in LUAD and LUSC (Kaplan-Meier plotter).

with 95% confidence intervals and log rank p value was calculated (p value < 0.05).

2.4. cBioPortal. cBioPortal (<http://www.cbioportal.org/>) is a comprehensive web resource that could visualize and analyze multidimensional cancer genomics data [28, 29]. In this study, we analyze the AT II-associated genes' multiple alterations for LUAD (TCGA, Pan-Cancer Atlas) and LUSC (TCGA, Pan-Cancer Atlas), which contained mutations, structural variants, and copy-number alterations.

2.5. STRING. STRING (<https://string-db.org/>) is a database of known and predicted protein-protein interactions (PPI) [30]. In this study, we conducted associations among the PPI network of AT II-associated genes to explore the role of AT II-related genes' coexpressed genes with STRING.

2.6. GeneMANIA. GeneMANIA (<http://www.genemania.org>) is a useful website that can find information on protein-protein, protein-DNA, and genetic interactions, pathways, reactions, gene and protein expression data, protein

levels of 8 AT II-associated genes in lung cancer and normal lung tissues using the ONCOMINE database. As the results shown in Figure 1, the expression levels of AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD were all remarkably reduced in lung cancer vs. normal tissues in multiple datasets. Furthermore, we compared the expressions of the eight AT II-associated genes in LUAD (483 LUAD and 347 normal tissues) and LUSC (486 LUSC and 338 normal tissues) by GEPIA. The results in Figure 2 indicated that the expression of AQP4, CLDN18, PGC, SFTPB, SFTPC, and SFTPD decreased in LUAD tissues and AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD decreased in in the LUSC tissues. We also contrasted the relative expression levels of eight AT II-related genes in LUAD and LUSC tissues and determined that among all the factors we evaluated, SFTPB was the highest expression in both LUAD and LUSC (Figure S1). Taken together, our results showed that the expressions of AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD were significant decrease in in NSCLC.

3.2. Correlation Between AT II-associated genes and Tumor Stages of NSCLC Patients. Lung cancer is divided into four stages according to the disease progression. As the condition develops, the patient's physiology and physical condition will also constantly change. Therefore, we assessed the correlation between the expression of AT II-associated genes and the patients' pathological cancer stages of LUAD and LUSC patients by using GEPIA. We found that the expression of all eight AT II-associated genes are significant correlated with pathological stage of NSCLC (Figure 3). AQP4 ($p = 1.81e - 06$), CLDN18 ($p = 4.64e - 06$), FOXA2 ($p = 1.28e - 04$), NKX2-1 ($p = 7.56e - 04$), PGC ($p = 3.08e - 07$), SFTPB ($p = 3.33e - 07$), SFTPC ($p = 1.4e - 08$), and SFTPD ($p = 1.54e - 07$) show that the AT II-associated genes were inclined to high expression in NSCLC patients with advanced cancer stages (Figure 3). These data suggested that the 8 AT II-associated might play a significant role in the tumorigenesis and progression of NSCLC.

3.3. Prognostic Features of AT II-associated genes in Patients with Lung Cancer. To analyze the prognostic values of AT II-associated genes in NSCLC patients, we assessed the correlation between these genes and overall survival (OS), progression-free survival (FP), and postprogression survival (PPS) using Kaplan-Meier plotter (Table 1). The results shown in Figure 4 and Figure S2 are the low-expression of genes including AQP4 (HR = 0.74, $p = 2.40e - 04$), CLDN18 (HR = 0.76, $p = 1.9e - 05$), FOXA2 (HR = 0.63, $p = 1.6e - 12$), NKX2-1 (HR = 0.67, $p = 4.9e - 10$), PGC (HR = 0.69, $p = 1e - 08$), SFTPB (HR = 0.67, $p = 6.3e - 10$), SFTPC (HR = 0.81, $p = 1.40e - 03$), and SFTPD (HR = 0.66, $p = 1.6e - 10$) were significantly associated with low OS. And the low-expression of CLDN18 (HR = 0.72, $p = 9.10e - 04$), FOXA2 (HR = 0.68, $p = 6.7e - 05$), NKX2-1 (HR = 0.81, $p = 3.10e - 02$), PGC (HR = 0.7, $p = 2.40e - 04$), SFTPB (HR = 0.82, $p = 4.80e - 02$), SFTPC (HR = 0.82, $p = 4.00e$

$- 02$), and SFTPD (HR = 0.68, $p = 6.2e - 05$) were significantly related to a reduced FP. Low-expression of CLDN18 (HR = 0.98, $p = 3.20e - 02$), FOXA2 (HR = 0.74, $p = 2.10e - 02$), and SFTPD (HR = 0.96, $p = 2.10e - 02$) apparently led to a short PPS. Moreover, no significant difference was found between the AT II-associated genes and disease-free survival (DFS) in NSCLC patients (Table 1).

3.4. Genetic Alteration and PPI Analyses of AT II-associated genes. Epigenetic alteration plays a vital role in early malignancies, so a comprehensive analysis of the molecular characteristics of AT II-associated genes was further performed in the LUAD and LUSC samples, respectively. We used the cBioPortal online tool to analyze the AT II-associated genes alterations in LUAD (TCGA, Pan-Cancer Atlas) and LUSC (TCGA, Pan-Cancer Atlas). The results demonstrated that the alterations of 8 AT II-associated genes covered 273 samples out of 1053 patients with NSCLC (26%) (Figure 5(a)). Moreover, the mutation rates of AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD were 3, 5, 2.4, 9, 2.8, 1.8, 5, and 1.1% of the investigated lung cancer samples, respectively (Figure 5(a)).

Moreover, a PPI network analysis of AT II-related genes was conducted with STRING. The results in Figure 5(b) illustrated that the DMBT1 gene which is a candidate tumor suppressor gene discovered in recent years was closely connected with AT II-associated genes (Figure 5(b)). Besides, some genes that play an important role in immune response regulation, blood cell proliferation, defense mechanisms, and acute phase response genes are also significantly connected with AT II-associated genes, including Microfibril-associated glycoprotein 4 (MFAP4), Pulmonary surfactant-associated protein A1 (SFTPA1) (Figure 5(b)). The GENEMANIA results also revealed the functions of the differentially expressed AT II-associated genes, which including Leucine-rich repeat kinase 2 (LRRK2), lysosomal-associated membrane protein 3 (LAMP3), Cathepsin E (CTSE0), ATP-binding cassette transporter A3 (ABCA3), forkhead box F1 (FOXF1), and Napsin A (NAPSA), and these genes were mainly related to lung development, late endosome, aspartic-type peptidase activity. (Figure 5(c)).

3.5. Immune Cell Infiltration of AT II-associated genes in Patients With NSCLC. Immune cell level is associated with the proliferation and progression of the cancer cell. In this study, to investigate the relationship between AT II-associated genes and cancer-related inflammation and the infiltration of immune cells, we use the TIMER to reveal a comprehensive analysis of the correlation between eight AT II-associated genes and immune cell infiltration (Figure S3 and S4). ALL AT II-associated genes (including AQP4, FOXA2, NKX2-1, PGC, SFTPB, SFTPD, CLDN18, and SFTPC) were positively associated with the infiltration of six immune cell types (CD8+ T cells, B cells, CD4+ T cells, macrophages, neutrophils, and dendritic cells; all $p < 0.05$) in LUSC and positively associated with the infiltration of B cells in LUAD ($p < 0.05$). SFTPC and CLDN18 were positive connection with the infiltration of B cells and six immune

cell types (CD8+ T cells, B cells, CD4+ T cells, macrophages, neutrophils, and dendritic cells) both in LUAD and LUSC cancers.

4. Discussion

The occurrence of lung cancer is a multistep process. For example, LUAD has always been thought to progress from atypical adenomatous hyperplasia (AAH) to adenocarcinoma in situ (AIS) [33]. Before the development of LUSC, we can observe preinvasive lesions in the airways [34]. Distinct molecular events and other malignant phenotypes make normal lung cells gain or lose certain functions leading to deregulation of key genetic signals involved in cell proliferation, differentiation, apoptosis, migration, and invasion [35–37]. Studies have shown that AT II cells can dedifferentiate into a cell stem-like state, which can continuously differentiate, proliferate, repair, and cause damage. Therefore, AT II is suspected to be the cell of origin in oncogene-driven lung cancers and can help maintain tumor progression [24].

In recent years, 8 AT II-associated genes have been confirmed to play key roles in tumor growth and development. For example, FOXA2 has been proved that it plays crucial roles in lung morphogenesis, surfactant protein production, goblet cell differentiation, and mucin expression [38]. Besides, Liu et al. experimentally found that the histone demethylase PHF8 can drive neuroendocrine prostate cancer (NEPC) development by epigenetically upregulation of FOXA2 [39]. Thyroid transcription factor 1 (TTF-1 or NKX2-1) has been known as an important development regulator of driving the brain, lungs, and thyroid maturation and morphogenesis [40]. Studies have demonstrated that NKX2-1 gene mutations related to compensated congenital hypothyroidism and unexplained respiratory distress due to lung hypoplasia in neonates [41]. NKX2-1 amplification and overexpression also have been proved to have contributed to lung cancer cell proliferation rates and survival results [42]. Interestingly, some researchers found an opposite phenomenon that NKX2-1 can constrain lung adenocarcinoma in part by repressing the embryonically restricted chromatin regulator Hmga2 [43]. Thus, the oncogenic and inhibitory function of NKX2-1 in the same tumor type confirms its role as a bifunctional lineage factor. Aquaporins (AQPs) are water channel proteins that can be capable of selectively transporting water and other small solutes across cells [44, 45]. In the lung, AQPs were supposed to facilitate fluid transportation in alveolar space, airway humidification, pleural fluid absorption, and submucosal gland secretion. AQP4 is one of members of the aquaporin family which was first discovered in 1994 [45, 46]. The change of AQP4 expression is associated with many central nervous system (CNS) diseases including epilepsy, edema, stroke, and glioblastoma [47]. Besides, in breast cancer, thyroid carcinoma (undifferentiated), and stomach cancer, the expression of AQP4 is low [48–51]. On the contrary, studies found that AQP4 is highly expressed in lung cancer and is involved in the invasion of lung cancer cells [52, 53]. Surfactant proteins (SP) are involved in surfactant function and innate immu-

nity in the human lung. In cystic fibrosis (CF), the genetic contribution of the surfactant protein genes, SFTPB, SFTPC, and SFTPD are contained [54]. Finally, CLDN18 is required for intercellular connectivity and has been reported to be involved in cell migration and metastasis, making it an oncogene in various cancer types, including pancreatic, esophageal, ovarian, and lung cancer [55].

In this study, we first systematically analyzed the expression of eight AT II-associated genes (AQP4, SFTPB, SFTPC, SFTPD, CLDN18, FOXA2, NKX2-1, and PGC) in lung cancer. The expression levels of the eight genes in lung cancer were lower. Additionally, we also verified that the expression of AT II-associated genes was observably related to clinical cancer stages in NSCLC patients. These results indicate that all these eight AT II-associated genes might take a significant part in the tumorigenesis and progression of NSCLC. Besides, all these eight AT II-associated genes were found to be notably related to OS in lung cancer patients, and low-expression was associated with short OS in lung cancer patients. Seven genes except AQP4 were significantly positive associated with FP. And low-expression of CLDN18, FOXA2, and SFTPD apparently led to a short PPS. All these results indicate that AT II-associated genes might be a protectable factor for survivals of NSCLC patients and thus might be potential prognostic biomarkers. In addition, our study showed that the expression level of AT II-associated genes was significantly correlated with the infiltration of six immune cell types. This result also suggests that AT II-associated genes may also reflect the immune status besides the disease prognosis.

5. Conclusion

In conclusion, this study provided clues of the values of AT II-associated genes (AQP4, SFTPB, SFTPC, SFTPD, CLDN18, FOXA2, NKX2-1, and PGC) as clinical biomarkers and therapeutic targets in NSCLC. We believe that these eight AT II-associated genes were expected to become new prognostic biomarkers in NSCLC and provide some new inspirations to assist in the design of new immunotherapies.

Data Availability

All data generated or analyzed during this study are included in this published article.

Disclosure

A preprint of this article has previously been published on Research Square [56].

Conflicts of Interest

The authors declared no conflicts of interest.

Authors' Contributions

L.P.R. and X.X.W. collected and analyzed the data. M.J.X.L. rewrote the manuscript. Y.X., P.L., and H.C.L. participated in discussions for this work. L.P.R., P.T., and L.X. directed

the project and wrote the manuscript. All named authors read and approved the final manuscript. Liping Ren and Xiaoxia Wen contributed equally to this work.

Acknowledgments

We thank the National Natural Science Foundation of China (81872957, 82130112).

Supplementary Materials

Supplementary 1. Figure S1. The relative expression level of AT II-associated genes in LUAD and LUSC (GEPIA). The darker the color of the bar, the higher the relative expression. The result evaluated that SFTPB was the highest expression in both LUAD and LUSC.

Supplementary 2. Figure S2. Prognostic value of AT II-associated genes (SFTPB, SFTPC, SFTPD, and PGC) in LUAD and LUSC (Kaplan-Meier plotter).

Supplementary 3. Figure S3. Correlations between AT II-associated genes and immune cell infiltration (TIMER). Correlations between the abundance of immune cells and the expression of AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD in LUAD.

Supplementary 4. Figure S4. Correlations between AT II-associated genes and immune cell infiltration (TIMER). Correlations between the abundance of immune cells and the expression of AQP4, CLDN18, FOXA2, NKX2-1, PGC, SFTPB, SFTPC, and SFTPD in LUSC.

References

- [1] W. Liang, Y. Zhao, W. Huang et al., “Non-invasive diagnosis of early-stage lung cancer using high-throughput targeted DNA methylation sequencing of circulating tumor DNA (ctDNA),” *Theranostics*, vol. 9, no. 7, pp. 2056–2070, 2019.
- [2] R. S. Herbst, D. Morgensztern, and C. Boshoff, “The biology and management of non-small cell lung cancer,” *Nature*, vol. 553, no. 7689, pp. 446–454, 2018.
- [3] J. Ferlay, M. Colombet, I. Soerjomataram et al., “Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods,” *International Journal of Cancer*, vol. 144, no. 8, pp. 1941–1953, 2019.
- [4] V. Relli, M. Trerotola, E. Guerra, and S. Alberti, “Abandoning the notion of non-small cell lung cancer,” *Trends in Molecular Medicine*, vol. 25, no. 7, pp. 585–594, 2019.
- [5] A. Roointan, T. Ahmad Mir, S. Ibrahim Wani et al., “Early detection of lung cancer biomarkers through biosensor technology: a review,” *Journal of Pharmaceutical and Biomedical Analysis*, vol. 164, pp. 93–103, 2019.
- [6] J. Didkowska, U. Wojciechowska, M. Mańczuk, and J. Łobaszewski, “Lung cancer epidemiology: contemporary and future challenges worldwide,” *Annals of Translational Medicine*, vol. 4, no. 8, p. 150, 2016.
- [7] Y. Huang, J. Wang, Y. Zhao et al., “cncRNadb: a manually curated resource of experimentally supported RNAs with both protein-coding and noncoding function,” *Nucleic Acids Research*, vol. 49, no. D1, pp. D65–d70, 2021.
- [8] C. Lin, H. Song, C. Huang et al., “Alveolar type II cells possess the capability of initiating lung tumor development,” *PLoS One*, vol. 7, no. 12, article e53817, 2012.
- [9] C. R. Sears and P. J. Mazzone, “Biomarkers in lung cancer,” *Clinics in Chest Medicine*, vol. 41, no. 1, pp. 115–127, 2020.
- [10] M. M. Clausen and S. W. Langer, “Improving the prognosis for lung cancer patients,” *Acta Oncologica*, vol. 58, no. 8, pp. 1077–1078, 2019.
- [11] T. Inage, T. Nakajima, I. Yoshino, and K. Yasufuku, “Early lung cancer detection,” *Clinics in Chest Medicine*, vol. 39, no. 1, pp. 45–55, 2018.
- [12] I. Toumazis, M. Bastani, S. S. Han, and S. K. Plevritis, “Risk-based lung cancer screening: a systematic review,” *Lung Cancer*, vol. 147, pp. 154–186, 2020.
- [13] Y. Zhang, T. Liu, J. Wang et al., “Cellinker: a platform of ligand-receptor interactions for intercellular communication analysis,” *Bioinformatics*, vol. 37, no. 14, pp. 2025–2032, 2021.
- [14] Z. M. Zhang, J. S. Wang, H. Zulfiqar, H. Lv, F. Y. Dao, and H. Lin, “Early diagnosis of pancreatic ductal adenocarcinoma by combining relative expression orderings with machine-learning method,” *Frontiers in Cell and Development Biology*, vol. 8, article 582864, 2020.
- [15] Z. M. Zhang, J. X. Tan, F. Wang, F. Y. Dao, Z. Y. Zhang, and H. Lin, “Early diagnosis of hepatocellular carcinoma using machine learning method,” *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 254, 2020.
- [16] Y. Zhang, T. Liu, L. Chen et al., “RIScoper: a tool for RNA-RNA interaction extraction from the literature,” *Bioinformatics*, vol. 35, no. 17, pp. 3199–3202, 2019.
- [17] M. Aspal and R. L. Zemans, “Mechanisms of ATII-to-ATI cell differentiation during lung regeneration,” *International Journal of Molecular Sciences*, vol. 21, no. 9, p. 3188, 2020.
- [18] R. J. Mason, “Biology of alveolar type II cells,” *Respirology*, vol. 11, no. s1, pp. S12–S15, 2006.
- [19] L. Zhao, M. Yee, and M. A. O’Reilly, “Transdifferentiation of alveolar epithelial type II to type I cells is controlled by opposing TGF- β and BMP signaling,” *American Journal of Physiology. Lung Cellular and Molecular Physiology*, vol. 305, no. 6, pp. L409–L418, 2013.
- [20] B. Ruaro, F. Salton, L. Braga et al., “The history and mystery of alveolar epithelial type II cells: focus on their physiologic and pathologic role in lung,” *International Journal of Molecular Sciences*, vol. 22, no. 5, p. 2566, 2021.
- [21] T. Rana, C. Jiang, G. Liu et al., “PAI-1 regulation of TGF- β 1-induced alveolar type II cell senescence, SASP secretion, and SASP-mediated activation of alveolar macrophages,” *American Journal of Respiratory Cell and Molecular Biology*, vol. 62, no. 3, pp. 319–330, 2020.
- [22] M. Selman and A. Pardo, “Revealing the pathogenic and aging-related mechanisms of the enigmatic idiopathic pulmonary fibrosis. an integral model,” *American Journal of Respiratory and Critical Care Medicine*, vol. 189, no. 10, pp. 1161–1172, 2014.
- [23] T. Shaurova, L. Zhang, D. W. Goodrich, and P. A. Hershberger, “Understanding lineage plasticity as a path to targeted therapy failure in EGFR-mutant non-small cell lung cancer,” *Frontiers in Genetics*, vol. 11, p. 281, 2020.
- [24] A. Maynard, C. E. McCoach, J. K. Rotow et al., “Therapy-induced evolution of human lung cancer revealed by single-

- cell RNA sequencing,” *Cell*, vol. 182, no. 5, pp. 1232–1251.e22, 2020.
- [25] D. R. Rhodes, J. Yu, K. Shanker et al., “ONCOMINE : a cancer microarray database and integrated data-mining platform,” *Neoplasia*, vol. 6, no. 1, pp. 1–6, 2004.
- [26] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, “GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses,” *Nucleic Acids Research*, vol. 45, no. W1, pp. W98–w102, 2017.
- [27] Á. Nagy, A. Lániczky, O. Menyhárt, and B. Györfly, “Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets,” *Scientific Reports*, vol. 8, no. 1, p. 9227, 2018.
- [28] E. Cerami, J. Gao, U. Dogrusoz et al., “The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data,” *Cancer Discovery*, vol. 2, no. 5, pp. 401–404, 2012.
- [29] J. Gao, B. A. Aksoy, U. Dogrusoz et al., “Integrative analysis of complex cancer genomics and clinical profiles using the cBio-Portal,” *Science Signaling*, vol. 6, no. 269, p. p11, 2013.
- [30] D. Szklarczyk, A. L. Gable, D. Lyon et al., “STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets,” *Nucleic Acids Research*, vol. 47, no. D1, pp. D607–d613, 2019.
- [31] D. Warde-Farley, S. L. Donaldson, O. Comes et al., “The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function,” *Nucleic Acids Research*, vol. 38, supplement 2, pp. W214–W220, 2010.
- [32] T. Li, J. Fan, B. Wang et al., “TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells,” *Cancer Research*, vol. 77, no. 21, pp. e108–e110, 2017.
- [33] K. Inamura, “Clinicopathological characteristics and mutations driving development of early lung adenocarcinoma: tumor initiation and progression,” *International Journal of Molecular Sciences*, vol. 19, no. 4, p. 1259, 2018.
- [34] A. Pennycuik, V. H. Teixeira, K. Abduljabbar et al., “Immune surveillance in clinical regression of preinvasive squamous cell lung cancer,” *Cancer Discovery*, vol. 10, no. 10, pp. 1489–1499, 2020.
- [35] F. Zito Marino, R. Bianco, M. Accardo et al., “Molecular heterogeneity in lung cancer: from mechanisms of origin to clinical implications,” *International Journal of Medical Sciences*, vol. 16, no. 7, pp. 981–989, 2019.
- [36] V. M. L. de Sousa and L. Carvalho, “Heterogeneity in lung cancer,” *Pathobiology*, vol. 85, no. 1-2, pp. 96–107, 2018.
- [37] Y. Zhang, T. Liu, X. Hu et al., “CellCall: integrating paired ligand-receptor and transcription factor activities for cell-cell communication,” *Nucleic Acids Research*, vol. 49, no. 15, pp. 8520–8534, 2021.
- [38] W. Choi, S. Choe, and G. W. Lau, “Inactivation of FOXA2 by respiratory bacterial pathogens and dysregulation of pulmonary mucus homeostasis,” *Frontiers in Immunology*, vol. 11, p. 515, 2020.
- [39] Q. Liu, J. Pang, L. A. Wang et al., “Histone demethylase PHF8 drives neuroendocrine prostate cancer progression by epigenetically upregulating FOXA2,” *The Journal of Pathology*, vol. 253, no. 1, pp. 106–118, 2021.
- [40] K. J. Peall, D. Lumsden, R. Kneen et al., “Benign hereditary chorea related to NKX2.1: expansion of the genotypic and phenotypic spectrum,” *Developmental Medicine and Child Neurology*, vol. 56, no. 7, pp. 642–648, 2014.
- [41] E. Kostopoulou, K. Miliordos, and B. Spiliotis, “Genetics of primary congenital hypothyroidism—a review,” *Hormones (Athens, Greece)*, vol. 20, no. 2, pp. 225–236, 2021.
- [42] K. A. Kwei, Y. H. Kim, L. Girard et al., “Genomic profiling identifies *TTF1* as a lineage-specific oncogene amplified in lung cancer,” *Oncogene*, vol. 27, no. 25, pp. 3635–3640, 2008.
- [43] M. M. Winslow, T. L. Dayton, R. G. Verhaak et al., “Suppression of lung adenocarcinoma progression by *Nkx2-1*,” *Nature*, vol. 473, no. 7345, pp. 101–104, 2011.
- [44] L. Ren, Y. Xu, L. Ning et al., “TCM2COVID: a resource of anti-COVID-19 traditional Chinese medicine with effects and mechanisms,” *iMETA*, p. e42, 2022.
- [45] D. He, A. Zhang, Y. Li, G. Cai, Y. Li, and S. Guo, “Autoimmune aquaporin-4 induced damage beyond the central nervous system,” *Multiple Sclerosis and Related Disorders*, vol. 18, pp. 41–46, 2017.
- [46] H. Hasegawa, T. Ma, W. Skach, M. A. Matthay, and A. S. Verkman, “Molecular cloning of a mercurial-insensitive water channel expressed in selected water-transporting tissues,” *The Journal of Biological Chemistry*, vol. 269, no. 8, pp. 5497–5500, 1994.
- [47] A. Vandebroek and M. Yasui, “Regulation of AQP4 in the central nervous system,” *International Journal of Molecular Sciences*, vol. 21, no. 5, p. 1603, 2020.
- [48] M. C. Papadopoulos and S. Saadoun, “Key roles of aquaporins in tumor biology,” *Key Roles of Aquaporins in Tumor Biology*, vol. 1848, no. 10 Part B, pp. 2576–2583, 2015.
- [49] D. Niu, T. Kondo, T. Nakazawa et al., “Differential expression of aquaporins and its diagnostic utility in thyroid cancer,” *PLoS One*, vol. 7, no. 7, article e40770, 2012.
- [50] Z. Shi, T. Zhang, L. Luo et al., “Aquaporins in human breast cancer: identification and involvement in carcinogenesis of breast cancer,” *Journal of Surgical Oncology*, vol. 106, no. 3, pp. 267–272, 2012.
- [51] H. Xu, Y. Zhang, W. Wei, L. Shen, and W. Wu, “Expression différentielle de l’aquaporine-4 dans la muqueuse gastrique normale et cancéreuse,” *Gastroentérologie Clinique et Biologique*, vol. 33, no. 1, pp. 72–76, 2009.
- [52] Y. Song, L. Wang, J. Wang, and C. Bai, “Aquaporins in respiratory system,” *Advances in Experimental Medicine and Biology*, vol. 969, pp. 115–122, 2017.
- [53] Y. Xie, X. Wen, Z. Jiang, H. Q. Fu, H. Han, and L. Dai, “Aquaporin 1 and aquaporin 4 are involved in invasion of lung cancer cells,” *Clinical Laboratory*, vol. 58, no. 1-2, pp. 75–80, 2012.
- [54] Z. Lin, N. Thorenoor, R. Wu et al., “Genetic association of pulmonary surfactant protein genes, SFTPA1, SFTPA2, SFTPB, SFTPC, and SFTPD with cystic fibrosis,” *Frontiers in Immunology*, vol. 9, p. 2256, 2018.
- [55] Y. L. Wan, H. J. Dai, W. Liu, and H. T. Ma, “miR-767-3p inhibits growth and migration of lung adenocarcinoma cells by regulating *CLDN18*,” *Oncology Research*, vol. 26, no. 4, pp. 637–644, 2018.
- [56] X. Wen, Y. Xiao, P. Leng, and H. Luo, *Comprehensive Analysis of Prognostic Value and Immune Infiltration of ATRX-associated Genes in Non-small Cell Lung Cancer*, PREPRINT (Version 1) available at Research Square, 2021.