Prognostic significance of peripheral blood S100A12, S100A8, and S100A9 concentrations in idiopathic pulmonary fibrosis

Dongyan Ding
The Second Hospital of Jilin University
Rumei Luan
The Second Hospital of Jilin University
Qianfei Xue
The University Hospital of Jilin University
Junling Yang (junling@jlu.edu.cn)
The Second Hospital of Jilin University

Research Article

Keywords: idiopathic pulmonary fibrosis (IPF), prognostic factors, S100A12, S100A8, S100A9, calgranulin

Posted Date: October 13th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3122211/v2

Additional Declarations: No competing interests reported.

Abstract

Background: S100A12, S100A8, and S100A9 are inflammatory disease biomarkers whose functional significance in idiopathic pulmonary fibrosis (IPF) remains unclear. We evaluated the significance of S100A12, S100A8, and S100A9 levels in IPF development and prognosis.

Methods: The dataset was collected from the Gene Expression Omnibus (GEO) database and differentially expressed genes were screened using GEO2R. We conducted a retrospective study of 106 patients with IPF to explore the relationships between different biomarkers and poor outcomes. Pearson's correlation coefficient, Kaplan–Meier, Cox regression, and functional enrichment analyses were used to evaluate relationships between these biomarkers' levels and clinical parameters or prognosis.

Results: Serum levels of S100A12, S100A8, and S100A9 were significantly elevated in patients with IPF. The two most significant co-expression genes of S100A12 were S100A8 and S100A9. Patients with levels of S100A12 (median 231.21 ng/mL), S100A9 (median 57.09 ng/mL) or S100A8 (median 52.20 ng/mL), as well as combined elevated S100A12, S100A9, and S100A8 levels, exhibited shorter progression-free survival and overall survival. Serum S100A12 and S100A8, S100A12 and S100A9, S100A9 and S100A8 concentrations also displayed a strong positive correlation (r_s^2 = 0.4558, r_s^2 = 0.4558, r_s^2 = 0.6373; P < 0.001). S100A12 and S100A8/9 concentrations were independent of FVC%, DLCO%, and other clinical parameters (age, laboratory test data, and smoking habit). Finally, in multivariate analysis, the serum levels of S100A12, S100A8, and S100A9 were significant prognostic factors (hazard ratio 1.002, P = 0.032, hazard ratio 1.039, P = 0.001, and hazard ratio 1.048, P = 0.003).

Conclusions: S100A12, S100A8, and S100A9 are promising circulating biomarkers that may aid in determining IPF patient prognosis. Multicenter clinical trials are needed to confirm their clinical value.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a progressive, life-threatening interstitial lung disease (1) associated with extracellular matrix remodeling, lipid metabolism, and autoimmune reactivity (2, 3). The overall 5-year survival rate for IPF is 20–40% (4). Treatment failure in IPF patients may be due, in part, to a limited understanding of IPF and a lack of predictive diagnostic/prognostic biomarkers. With the successful introduction of the anti-fibrosis drugs, pirfenidone, and nintedanib, in recent years, anti-IPF therapy has been demonstrated to have potential (5), and early intervention may help improve clinical outcomes (2).

Calgranulins (S100A8, S100A9, and S100A12) are members of the S100 small calcium-binding protein family. They are damage-associated molecular pattern (DAMP) molecules that serve as important inflammatory mediators involved in many intracellular processes, including cell proliferation and differentiation, and calcium homeostasis (6). S100A12 and S100A8/9 stimulate neutrophil and monocyte chemotaxis, fibrin adherence *in vitro*, neutrophil migration to sites of inflammation *in vivo* (7, 8), binding and activation of Toll-like receptor 4 (TLR4) and advanced glycosylation end-product receptor

(RAGE), and altering inflammation by promoting cellular immune responses (9, 10). These three S100 proteins also activate airway epithelial cells to produce mucin 5AC (11), and inhibit lung fibroblast migration (12). Previous studies have suggested that S100A12 and S100A8/A9 may be closely associated with IPF(10), but to date, the clinical relevance of serum S100A12 and S100A8/9 concentrations in patients with IPF remains to be investigated. We thus hypothesized that serum S100A12 and S100A8/9 concentrations in patients with IPF reflect lung fibroblast activity and serve as useful prognostic biomarkers. Due to the lack of functional S100A12 in mice (13), we compared S100A12 concentrations with S100A8/A9 concentrations in patients with IPF.

We used the Gene Expression Omnibus (GEO) database to determine the differential concentrations of S100A12 and S100A8/9 in patients with IPF. Associations between clinical parameters and mortality were evaluated by further quantifying serum S100A12 and S100A8/9 concentrations in patients with IPF. This study provides insights into the molecular mechanisms underlying IPF and identifies potential molecular targets for the development of novel intervention strategies.

Materials And Methods

2.1. Gene expression data

The gene expression dataset GSE93606, from the National Center for Biotechnology Information Comprehensive Gene Expression Database (https://www.ncbi.nlm.nih.gov/geo/) was used. GSE93606 includes the expression profiles of 60 peripheral blood samples from patients with IPF and 20 samples from healthy donors. GSE93606 data were generated using an Affymetrix Human Gene 1.1 ST Array (Affymetrix, Inc., Santa Clara, CA, USA; GPL11532 platform).

2.2. Differentially expressed gene data preprocessing and screening

GEO2R is a web-based tool within the GEO database specifically designed to analyze differentially expressed genes (DEGs) between experimental and control groups. GEO2R analysis was performed to compare IPF and control groups, and the differential expression results were downloaded for further analysis. The critical criteria for differential gene expression were an absolute value of logarithmic fold change (FC) > 1 and a P value < 0.05.

2.3. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis of DEGs

Gene Ontology (GO) enrichment of DEGs in the biological process, molecular function, and cellular component categories, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the DAVID online tool (https://david.ncifcrf.gov and https://bioinformatics.com.cn/) (14). Forced vital capacity (predicted

FVC%), diffusion capacity of the lung for carbon monoxide (predicted DLCO%), progression-free survival (PFS), and gender-age-physiology (GAP) index value were calculated for the GSE93606 dataset based on previous reports. PFS was defined as disease progression (defined as FVC% decline predicted within 6 months > 10%) at follow-up, death, or review at the time of last contact.

2.4. Construction of a protein-protein interaction network

The STRING database analyzes protein-protein interactions (PPIs) using online tools (http://www.string.embl.de/, version:11.5) (15). PPI analysis was carried out using the STRING database, with a confidence score > 0.4 as the cutoff. The Multi-Experiment Matrix (MEM) database (https://biit.cs.ut.ee/mem/index.cgi) was used to verify correlations between genes from the STRING database, using hundreds of public S100A12 gene expression datasets (16). Correlations between S100A12, S100A8, and S100A9 concentrations and lung function were determined using the Lianchuan Biocloud platform (OMicstudione.cn).

2.5. Study population

This prospective cohort study included patients with IPF who were hospitalized or were outpatients at the Second Hospital of Jilin University, the People's Hospital of Jilin Province, and the People's Hospital of Jilin City from June 2017 to June 2022 and who had serum samples collected at the time of diagnosis. The inclusion criteria included: 1) patients diagnosed with IPF based on the ATS/ERS/JRS/ALAT IPF Diagnostic and Management Guidelines (17) in a fashion of multidisciplinary discussion; 2) All patients underwent high-resolution computed tomography (HRCT) at diagnosis and images were independently reviewed by two radiologists who were blinded to the clinical information. And the exclusion criteria included: 1) Patients with other known causes of interstitial lung disease, such as connective tissue disease with autoimmune features, domestic or occupational environmental exposure, and drug toxicity, were excluded; 2) Patients with lung cancer, liver and kidney dysfunction, infection(tuberculosis, bacteria, mycology, etc), and lung transplantation affecting the level of serum S100A12, S100A8, and S100A9 are excluded from the cohort; 3) Those who refused consent or were lost to follow-up were also excluded. Based on these criteria, 106 patients with IPF were included in this study. Serum samples from 60 age-sex-matched healthy volunteers were collected as a control group of which all participants underwent a full medical examination prior to inclusion in the study. All patients were followed up for > 1 year or until death.

2.6. Data collection

Basic and clinical information (sex, age, laboratory test data, pulmonary function test results, and GAP index value) were collected from all patients with IPF at the time of enrollment, and venous blood samples were collected. All patients in our study were treated according to standard clinical practice guidelines. A follow-up investigation was conducted with enrolled participants by telephone until July

2023. Study participants were followed up every 3–4 months by the Simmons Center's customary care practices. Physiological data (lung function test results) and physician assessments were performed at all visits. All included patients were followed up clinically for at least 12 months (except for those who died). These studies were conducted according to the principles of the Declaration of Helsinki and were approved by the Human Investigation Committee of the Second Hospital of Jilin University (No. 2023111). All patients signed an informed consent form to participate in these studies.

2.7. Immunoassays

Peripheral blood samples collected from IPF patients were centrifuged at 3,000 × g for 10 min within 30 min of collection, and the resulting serum samples were stored at -80 °C for subsequent assays. Serum S100A12, S100A8, and S100A9 concentrations were determined using a commercially available enzyme-linked immunosorbent assay (ELISA) (Shanghai Enzyme Linked Biotechnology Co., Ltd., Shanghai, China), performed in batches by the same technician, who was blinded to the clinical data.

2.8. Statistical analyses

Statistical analyses were performed using SPSS software (version 20.0; SPSS Inc., Chicago, IL, USA) and GraphPad Prism software (version 5.0; GraphPad Software Inc., San Diego, CA, USA). Categorical variables are expressed as counts (percentages), and differences between groups were compared using chi-square tests or Fisher's precision tests, as appropriate. Normally distributed variables were presented as mean and SD, while non-normally distributed variables were presented as median and interquartile range (IQR). A Student's t-test and a Mann-Whitney U test were used for the between-group comparison. Bivariate correlations were analyzed using Spearman's correlation coefficients. The area under the curve and 95% confidence intervals (CIs) generated by the receiver operating characteristic (ROC) curve analysis were used to evaluate the predicted values. The Kaplan–Meier method was used to summarize the survival probability of IPF patients, and univariate and multivariate Cox proportional hazard regression models were used to determine the impact of variables on the survival of IPF patients. The associated hazard ratios (HRs) and 95% CIs were reported. $P \leq 0.05$ was considered statistically significant.

Results

3.1. Differential expression analysis of peripheral blood serum in patients with IPF

In the dataset, GSE93606, the DEGs of the IPF group and control group were analyzed by GEO2R, and the results were shown in Figure 1A and the UMAP cluster in Figure 1B. In this study, we have identified a total of 1,358 DEGs between serum samples from the two groups. Significant differences in gene expression distribution were observed between the two groups. The five tops differentially expressed

genes were thioredoxin (*TXN*; FC, 2.2; $P = 2.75 \times 10^{-9}$), cystatin A (*CSTA*; FC, 2.1; $P = 5.81 \times 10^{-9}$), chemokine-like factor superfamily member 2 (*CMTM2*; FC, 1.7; $P = 1.78 \times 10^{-8}$), *S100A12* (FC, 1.94; $P = 8.75 \times 10^{-8}$), and retinol-binding protein 7 (*RBP7*; FC, 1.86; $P = 1.45 \times 10^{-6}$). To reveal the potential biological functions and pathways associated with these DEGs, GO enrichment and KEGG pathway analyses were performed. According to GO and KEGG analyses (Fig. 1D), the DEGs were significantly correlated with positive regulation of the inflammatory response, extracellular exosome, the innate immune response, TLR4 binding, RAGE binding, and oxidative phosphorylation signaling pathways.

3.2. Expression and co-expression of S100A12 in peripheral blood serum of IPF patients

S100A12 expression was significantly upregulated in blood samples from patients with IPF, especially in those with a poor prognosis (Fig. 2C). According to the STRING database, the expression of 10 protein-coding genes (*S100A8*, *S100A9*, *IDH1*, *ANXA5*, *GAPDH*, *ALDOA*, *RELA*, *NFKB2*, *NFKB1*, and *AGER*) was regulated with *S100A12* (Fig. 1E). Co-expression scores based on RNA expression patterns, and on protein co-regulation provided by Proteome HD (Fig. 1F). S100A12 and S100A9 RNA co-expression score 0.927; S100A12 and S100A8 RNA co-expression score 0.983; S100A9 and S100A8 combined co-expression score 0.995 based on RNA expression 0.954 and protein co-regulation. Based on the MEM database, *S100A8* and *S100A9* were the most significantly co-regulated genes with *S100A12* (Fig. 1C). These genes were thus included in further studies. Both S100A8 and S100A9 were upregulated in blood samples from patients with IPF, especially in those with poor prognosis (Fig. 2A, B).

3.3. Correlation analysis of lung function and prognosis

No significant association was observed between serum S100A12, S100A8, and S100A9 levels and lung function (Fig. 1G). A ROC curve analysis was performed to determine the optimal threshold for predicting disease progression (Fig. 2G). Patients with IPF were divided into two groups based on the optimal cut-off value of gene expression levels. S100A12, S100A8, and S100A9 concentrations were significantly negatively associated with PFS (Figs. 2D, E, F). The calculated pooled values of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of S100A8 were 98%, 70%, 90%, and 93%, respectively. The calculated pooled values of sensitivity, specificity, PPV, NPV of S100A9 were 78%, 95%, 98%, and 59%, respectively. The calculated pooled values of sensitivity, specificity, PPV, NPV of S100A12 were 88%, 85%, 95%, and 68%, respectively.

3.4. Baseline characteristics of study participants

Study population characteristics are summarized in Table 1. No significant differences in age, sex, or body mass index (BMI) were observed between IPF patients and healthy controls. We initially evaluated 125 IPF patients and excluded twelve due to secondary causes (eight), autoimmune disease (six), and

malignancy (five). Finally, 106 patients with IPF and 60 healthy subjects were evaluated. Baseline pulmonary function test data in patients with IPF showed moderate pulmonary impairment, with a mean predicted FVC% of 75.98 ±14.14 and a mean predicted DLCO% of 52.18 ±14.84.

Table 1. Baseline characteristics of the study population.

	Controls	IPF	*P*-value
Number of subjects	60	106	
Gender			
Male, n (%)	32 53.33	55 51.89	0.858
Female, n (%)	28 46.67	51 48.11	
Age (years), mean (SD)	64.63 ± 3.29	65.62 ± 8.45	0.286
BMI, mean (SD)	22.54 ± 0.93	22.82 ± 1.24	0.099
Smoking status			
- never smokers, n (%)	38 (63.33)	60 (56.60)	0.397
- ex-smokers, n (%)	6 (10.00)	10 (9.43)	0.905
- current smokers, n (%)	16 (26.67)	36 (33.97)	0.330
WBC $*10^9$/L	7.93 ± 0.71	8.40 ± 2.31	0.052
CRP mg/dL	0.49 ± 0.22	0.44 ± 0.19	0.211
PCT ng/ml	0.29 ± 0.09	0.27 ± 0.08	0.180
KL6 ng/ml	1.35 ± 0.09	1.84 ± 0.29	<0.001
S100A8 ng/ml	19.50 18.61 23.46	52.70 28.76 72.52	<0.001
S100A9 ng/ml	14.46 13.32 16.24	56.16 27.98 69.19	<0.001
S100A12 ng/ml	91.97 84.99 102.83	226.21 161.36 522.39	<0.001
FVC (% of predicted), mean (SD)	N/A	75.98 ±14.14	
DLCO (% of predicted), mean (SD)	N/A	52.18 ±14.84	
FEV1 (% of predicted), mean (SD)	N/A	72.12 ±17.02	
GAP index, n (%)			

I	N/A	68 64.15
II	N/A	35 33.02
III	N/A	3 2.83

Abbreviations: IPF, idiopathic pulmonary fibrosis; SD standard deviation; WBC, white blood cell; CRP, C-reactive protein; PCT, procalcitonin; KL6, Krebs von den Lungen-6; FVC, forced vital capacity; DLCO diffusion capacity for carbon monoxide; GAP index Gender-Age-Physiology index; BMI Body Mass Index.

Among IPF patients, forty-three subjects were treated with pirfenidone (40.6%) and thirty-two with nintedanib (30.2%); thirty-one patients refused any antifibrotic drug (29.2%). There was no significant difference in terms of survival between patients treated with pirfenidone or nintedanib.

3.5. Serum S100A12 and S100A8/9 concentrations in IPF patients

Serum S100A12 and S100A8/9 concentrations were measured in all study subjects. We noted that the serum concentrations of S100A12, S100A8, and S100A9 in the control group and the IPF group were 91.97 (84.99, 102.83) ng/ml vs 226.21 (161.36 522.39) ng/ml, P<0.001, 19.50 (18.61, 23.46) ng/ml vs 52.70 (28.76, 72.52) ng/ml, P<0.001, 14.46 (13.32, 16.24) ng/ml vs 56.16 (27.98, 69.19) ng/ml, P<0.001. Compared to the control group, the concentration of IPF patients increased significantly (Fig. 3A, B, C).

3.6. S100A12 and S100A8/9 concentrations were associated with clinical parameters

Correlations between serum S100A12 concentration and several clinical parameters are shown in Figure 4. Serum S100A12 and S100A8 concentrations showed a strong positive correlation (Fig. 4C; r_s^2= 0.4558, P< 0.001). A strong positive correlation between serum S100A12 and S100A9 concentrations was also observed (Fig. 4B; r_s^2 = 0.5355, P< 0.001). Serum S100A9 and S100A8 concentrations also displayed a strong positive correlation (Fig. 4A; r_s^2 = 0.6373, P< 0.001). Serum S100A12, S100A8, and S100A9 concentrations in IPF patients correlated positively with the KL6 concentration (Fig. 4F; r_s^2 = 0.03481, P = 0.0555; Fig. 4D; r_s^2 = 0.02641, P = 0.0960; Fig. 4E; r_s^2 = 0.04754, P = 0.0247, respectively). S100A12 and S100A8/9 concentrations were independent of FVC%, DLCO%, and other clinical parameters (age, laboratory test data, and smoking habit).

3.7. Subgroup analysis based on S100A12 and S100A8/9 concentrations

A ROC curve analysis was performed to determine the optimal threshold for predicting disease progression (Fig. 3G, Table 2). With 116.2 ng/mL as the threshold serum S100A12 concentration,

sensitivity, specificity, PPV, NPV were 95%, 98%, 99%, and 92%, respectively. With 27.99 ng/mL as the threshold serum S100A8 concentration, sensitivity, specificity, PPV, NPV were 77%, 100%, 100%, and 71%, respectively. With 17.95 ng/mL as the threshold serum S100A9 concentration, sensitivity, specificity, PPV, NPV of S100A9 were 98%, 90%, 95%, and 96%, respectively.

Table 2. Discriminating capability and cut-off values of baseline serum S100A12、S100A8、and S100A9 by ROC curve analysis distinguishing IPF from control.

	S100A12	S100A8	S100A9
AUC	0.9912	0.9374	0.9857
95% CI	0.9811 -1.000	0.9038 - 0.9711	0.9730- 0.9984
P-value	< 0.001	< 0.001	< 0.001
Cut-off value	116.2 ng/mL	27.99 ng/mL	17.95 ng/mL
Sensitivity	95%	77%	98%
Specificity	98%	100%	90%
PPV	99%	100%	95%
NPV	92%	71%	96%

Abbreviations: S100A12, S100 calcium binding protein A12; S100A8, S100 calcium binding protein A8; S100A9, S100 calcium binding protein A9; AUC, area under the curve; CI, confidence interval.

3.8. Overall survival analysis

Of 106 patients with IPF, 62 (58.5%) exhibited disease progression. According to Kaplan–Meier analysis, IPF patients with high serum concentrations of S100A12 and S100A8/9 had poorer overall outcomes than healthy subjects and IPF patients with low serum concentrations of S100A12 and S100A8/9 (Fig. 3D, E, F). When patients were divided into two groups based on median S100A12 (231.21 ng/mL), median S100A8 (52.20 ng/mL), or median S100A9 (57.09 ng/mL), patients with higher levels had significantly lower survival rates than those with lower levels.

3.9. Prognostic factors for survival among patients with IPF

In univariate Cox proportional risk analysis, higher serum S100A12 and S100A8/9 concentrations were significant risk factors in patients with IPF (HR = 1.005, $P < 0.001$; HR = 1.059, $P < 0.001$; HR = 1.078; $P < 0.001$; Table 3). Higher KL6 concentrations were significantly associated with poorer survival (HR = 2.806, $P = 0.023$; Table 3). In multivariate analyses, only serum S100A12 and S100A8/9 concentrations (HR = 1.002, $P = 0.032$; HR = 1.039, $P = 0.001$; HR = 1.048, $P = 0.003$, respectively) were associated with mortality in patients with IPF (Table 3) and were independent predictors of higher rates of disease progression.

Table 3. Results of the Cox proportional hazards regression analysis of disease progression.

ɛters	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
	1.019 (0.988–1.051)	0.230	1.032 (0.966–1.103)	0.345
	0.868 (0.524–1.438)	0.583	1.431 (0.644–3.181)	0.379
	0.877 (0.705–1.091)	0.238	0.760 (0.554–1.042)	0.089
S100A12 ng/ml	1.005(1.003–1.006)	0.001	1.002(1.000–1.005)	0.032
S100A8 ng/ml	1.059 (1.043–1.075)	0.001	1.039 (1.015–1.064)	0.001
S100A9 ng/ml	1.078 (1.051–1.106)	0.001	1.048 (1.016–1.081)	0.003
10^9/L	0.961 (0.857–1.078)	0.498	1.031 (0.898–1.183)	0.665
g/dL	1.708 (0.466–6.262)	0.420	2.399 (0.381–15.124)	0.351
ŋ/ml	0.082(0.003–2.068)	0.129	0.023 (0.000–1.221)	0.063
ŋ/ml	2.806 (1.156–6.811)	0.023	1.527 (0.656–3.554)	0.327
ɾedicted	1.000 (0.983–1.017)	0.997	0.964 (0.928–1.001)	0.055
predicted	1.010 (0.992–1.027)	0.275	0.990 (0.959–1.022)	0.538
ɔredicted	0.998 (0.983–1.014)	0.837	1.003 (0.981–1.027)	0.775
dex, n (%)	0.939 (0.782–1.129)	0.506	0.740 (0.451–1.214)	0.234

Significance tests for comparisons between groups were based on Cox proportional hazards regression.

Abbreviations: IPF, idiopathic pulmonary fibrosis; WBC, white blood cell; CRP, C-reactive protein; PCT, procalcitonin; KL6, Krebs von den Lungen-6; FVC, forced vital capacity; DLCO diffusion capacity for carbon monoxide; GAP index Gender-Age-Physiology index; BMI Body Mass Index.

Discussion

We aimed to evaluate serum S100A12 and S100A8/9 concentrations and their prognostic significance in patients with IPF. Previous studies have demonstrated an association between inflammatory cytokines and innate and adaptive immune cell infiltration in IPF (18). S100A8, S100A9, and S100A12 are danger-signaling members of the DAMP family that initiate and amplify local inflammation and innate immune responses (19). For several inflammatory diseases, including inflammatory bowel disease, various rheumatic diseases, fever syndrome, and vasculitis, these three S100 proteins are considered typical alarm proteins (20). The peripheral blood expression profile has undergone corresponding modifications in the current investigation, which suggests that the existence of an altered or overactive inflammatory immune response is met with a host response. These responses persisted at elevated levels during long-term monitoring and varied across patients with stable and progressing disease, pointing to the possibility that an overactive inflammatory immune response may serve as a constant stimulus for recurrent alveolar injury in IPF.

In this investigation, we found 1358 DEGs in total between serum samples from IPF patients and healthy controls. These DEGs are closely linked to ECM-related functions, according to GO enrichment analysis, which is consistent with the clinical traits of IPF. This correlation corresponds to the clinical features of IPF, where the accumulation of ECM and declining parenchymal cells are distinct histological attributes driving abnormal lung remodeling and disease development (4, 21). Additionally, the analysis of KEGG pathways indicates the involvement of these differentially expressed genes (DEGs) in Spliceosome and oxidative phosphorylation signaling pathways. A previous study has demonstrated that Spliceosome functions as a proinflammatory DAMP in lung fibrosis (22, 23). Both of these pathways play critical roles in the pathogenesis of IPF (24).

Bioinformatics studies have shown that S100A12, S100A8, and S100A9 are upregulated in the blood of IPF patients relative to levels in healthy subjects. Levels are further elevated in IPF patients with a poor prognosis. Our ELISA results support these findings. The genes S100A9, S100A8, and S100A12 play a role in three biological processes associated with IPF, specifically ECM remodeling, lipid metabolism, and immune response (3). S100A8, S100A9, and S100A12 are mainly released by activated neutrophils and monocytes (7, 8), involved in activating signal transduction pathways in endothelial, vascular smooth muscle, and inflammatory cells, leading to the transcription and secretion of pro-inflammatory cytokines and cell adhesion molecules (25).

S100A12 inhibits migration of lung fibroblasts via RAGE-p38 MAPK signaling (12). A study has proposed a potential role for RAGE signaling in the development of abnormal lung tissue repair processes due to changes in mesenchymal cell migration (12). Thus, the effect of S100A12 on lung fibroblast migration might be caused by random migration (chemokinesis) and/or directed migration (chemotaxis) (12, 26). In addition, actin polymerization is crucial for cell migration (26), but the inhibition of migration by S100A12 was not caused by impaired actin polymerization. High concentrations of S100A12 in the peripheral blood predict a low overall survival rate for IPF, suggesting that an excessive inflammatory immune response worsens prognosis (7), S100A12 as Biomarker of Disease Severity and Prognosis in Patients With Idiopathic Pulmonary Fibrosis (27) consistent with the results of this study.

S100A8 and S100A9 are the most important co-expression genes of S100A12, once released from the cell, they act as DAMPs via TLR4 and RAGE (19). S100A8/A9 have bipolar functions. By stimulating the production of reactive oxygen species (ROS), S100A8/A9 induce the secretion of several pro-inflammatory cytokines, including IL-6, TNFα, and IL-1β. This in turn activates the transcription factor NF-κB, leading to cytokine secretion and expressing and activating the NLRP3 inflammasome (28). Blocking S100A8/A9 activity using small molecule inhibitors or antibodies ameliorates pathological conditions in mouse models by eliminating inflammatory activity. These findings suggest that S100A8/A9 has a pro-inflammatory nature (28, 29). In contrast, S100A8/A9 also have anti-inflammatory functions in various microenvironments. S100A8/A9 inhibit the differentiation and functioning of dendritic cells (30), induce myeloid-derived suppressor cells (31), and increase the functions of regulatory T-lymphocytes (32). In our study, we found a correlation between elevated levels of S100A8/A9 and decreased survival rates in patients suffering from IPF. The pro-inflammatory effects and subsequent anti-inflammatory effects of S100A8/A9 may contribute to compromised tissue repair and the advancement of lung fibrosis, similar to M1 and M2 macrophages (33). In terms of fibroblasts, which have a crucial role in fibrosis, S100A8/A9 activate cardiac fibroblasts (34) and dermal fibroblasts (35). These studies indicated that S100A8/A9 are related to the progression of fibrosis. Furthermore, GO and KEGG analyses showed that DEGs were associated with TLR4 binding and RAGE binding signaling pathways. These results are consistent with those of previous studies.

Multivariate analyses showed that higher serum S100A12 and S100A8/9 concentrations were associated independently with higher rates of disease progression, regardless of age, sex, BMI, or other variables. It can be speculated that greater dysregulation of inflammatory and fibrotic homeostasis in lung tissue amplifies lung injury, thus worsening patient outcomes (33). Our ROC and other results suggest that the effects of S100A8/A9 and S100A12 are highly similar. KL-6 is a marker of alveolar epithelial cell injury (36). Our univariate analysis revealed a significant correlation between KL-6 concentration and IPF prognosis, but multivariate analysis showed no significant correlation. This may be due to the gradual depletion of KL-6 producing cells, as normal lung area decreases with the progression of chronic IPF, resulting in lower serum KL-6 concentrations (37). We also observed higher peripheral blood white blood cell counts in patients with IPF compared with healthy subjects, and serum C-reactive protein concentrations tended to be higher, but correlations with IPF patient prognosis were not obvious. In the multivariate model generated from our derived cohort and tested in our validation cohort, S100A12 and

S100A8/9 together explained the overlapping proportion of poor prognostic variability in IPF. However, the protein concentration threshold from the derived cohort was not optimal for the validation cohort in univariate analysis. Our results thus indicate that combining molecular information with clinical parameters is important for deriving repeatable and accurate outcome prediction rules. Based on these results, serum S100A12 and S100A8/9 concentrations appear to be prognostic markers in patients with IPF.

Given that traditional physiological measures do not accurately predict short- and long-term prognoses, incorporating peripheral blood protein concentrations into IPF prognosis is a key step toward improving the classification and management of these patients. Our study was a prospective, multicenter study based on a well-defined group of patients with IPF. Study limitations include a lack of cytokine measurements in bronchoalveolar lavage fluid and tissues parallel to our measurements in peripheral blood. Finally, a larger study is needed to determine the prevalence of high serum S100A12 and S100A8/9 concentrations among IPF in patients and the optimal cut-offs for predicted outcomes. Future studies should repeatedly measure serum S100A12 and S100A8/9 concentrations to reveal predictors of disease progression or risk of death.

Conclusions

Important findings of this study include: 1) serum S100A12 and S100A8/9 concentrations were significantly increased in IPF patients and 2) higher serum S100A12 and S100A8/9 concentrations were associated with adverse outcomes in IPF patients. Novel features of this study include bioinformatic analysis of DEGs associated with IPF and validation of serum S100A12 and S100A8/9 concentrations by ELISA to predict IPF outcome. We recommend routine testing of serum S100A12 and S100A8/9 concentrations to stratify patients with worse prognoses.

Abbreviations

CI, confidence interval; DAMP, damage-associated molecular pattern; DLCO, diffusion capacity of the lung for carbon monoxide; FVC forced vital capacity; GAP, gender-age-physiology; GEO, Gene Expression Omnibus; GO, Gene Ontology; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; KEGG, Kyoto Encyclopedia of Genes and Genomes; NPV negative predictive value; PFS, progression-free survival; PPI, protein-protein interaction; PPV, positive predictive value; RAGE, advanced glycosylation end-product receptor; ROC, receiver operating characteristic; TLR4, Toll-like receptor 4.

Declarations

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics approval and consent to participate

This study was approved by the Human Investigation Committee of the Second Hospital of Jilin University (No. 2023111). All patients signed an informed consent form to participate in the study.

Consent for publication

Not applicable

Availability of data and materials

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

https:www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE93606

Funding

This work was supported by the National Key Technologies Research and Development Program (Grant No. 2021YFC2500700).

Authors' contributions

Dong-Yan Ding: Writing – original draft, Validation, Resources, Formal analysis; Jun-Ling Yang: Conceptualization, Writing – review & editing, Funding acquisition, Project administration; Ru-Mei Luan: Methodology, Data Curation; Qian-Fei Xue: Investigation. All authors have read and agreed to the published version of the manuscript.

Acknowledgments

We thank Editage Ltd. for the editorial work.

References

1. Spagnolo P, Kropski JA, Jones MG, Lee JS, Rossi G, Karampitsakos T, et al. Idiopathic pulmonary fibrosis: Disease mechanisms and drug development. Pharmacol Ther. 2021;222:107798 doi.org/10.1016/j.pharmthera.2020.107798.

2. Ma H, Wu X, Li Y, Xia Y. Research Progress in the Molecular Mechanisms, Therapeutic Targets, and Drug Development of Idiopathic Pulmonary Fibrosis. Front Pharmacol. 2022;13:963054 doi.org/10.3389/fphar.2022.963054.

3. Qian W, Xia S, Yang X, Yu J, Guo B, Lin Z, et al. Complex Involvement of the Extracellular Matrix, Immune Effect, and Lipid Metabolism in the Development of Idiopathic Pulmonary Fibrosis. Front Mol Biosci. 2021;8:800747 doi.org/10.3389/fmolb.2021.800747.

4. Richeldi L, Collard HR, Jones MG. Idiopathic pulmonary fibrosis. Lancet. 2017;389(10082):1941-52 doi.org/10.1016/s0140-6736(17)30866-8.

5. Finnerty JP, Ponnuswamy A, Dutta P, Abdelaziz A, Kamil H. Efficacy of antifibrotic drugs, nintedanib and pirfenidone, in treatment of progressive pulmonary fibrosis in both idiopathic pulmonary fibrosis (IPF) and non-IPF: a systematic review and meta-analysis. BMC Pulm Med. 2021;21(1):411 doi.org/10.1186/s12890-021-01783-1.

6. Gonzalez LL, Garrie K, Turner MD. Role of S100 proteins in health and disease. Biochim Biophys Acta Mol Cell Res. 2020;1867(6):118677 doi.org/10.1016/j.bbamcr.2020.118677.

7. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, et al. Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis. Am J Respir Crit Care Med. 2012;185(1):67-76 doi.org/10.1164/rccm.201101-0058OC.

8. Tanaka K, Enomoto N, Hozumi H, Isayama T, Naoi H, Aono Y, et al. Serum S100A8 and S100A9 as prognostic biomarkers in acute exacerbation of idiopathic pulmonary fibrosis. Respir Investig. 2021;59(6):827-36 doi.org/10.1016/j.resinv.2021.05.008.

9. Pietzsch J, Hoppmann S. Human S100A12: a novel key player in inflammation? Amino Acids. 2009;36(3):381-9 doi.org/10.1007/s00726-008-0097-7.

10. Zheng J, Dong H, Zhang T, Ning J, Xu Y, Cai C. Development and Validation of a Novel Gene Signature for Predicting the Prognosis of Idiopathic Pulmonary Fibrosis Based on Three Epithelial-Mesenchymal Transition and Immune-Related Genes. Front Genet. 2022;13:865052 doi.org/10.3389/fgene.2022.865052.

11. Kubo F, Ariestanti DM, Oki S, Fukuzawa T, Demizu R, Sato T, et al. Loss of the adhesion G-protein coupled receptor ADGRF5 in mice induces airway inflammation and the expression of CCL2 in lung endothelial cells. Respir Res. 2019;20(1):11 doi.org/10.1186/s12931-019-0973-6.

12. Tanaka N, Ikari J, Anazawa R, Suzuki M, Katsumata Y, Shimada A, et al. S100A12 inhibits fibroblast migration via the receptor for advanced glycation end products and p38 MAPK signaling. In Vitro Cell Dev Biol Anim. 2019;55(8):656-64 doi.org/10.1007/s11626-019-00384-x.

13. Bagheri V. S100A12 is not expressed in rodents: Transgenic mouse model is needed. Molecular Immunology. 2022;152:35 doi.org/https://doi.org/10.1016/j.molimm.2022.10.003.

14. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. 2022;50(W1):W216-w21 doi.org/10.1093/nar/gkac194.

15. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res. 2017;45(D1):D362-d8 doi.org/10.1093/nar/gkw937.

16. Kolde R, Laur S, Adler P, Vilo J. Robust rank aggregation for gene list integration and meta-analysis. Bioinformatics. 2012;28(4):573-80 doi.org/10.1093/bioinformatics/btr709.

17. Raghu G, Remy-Jardin M, Richeldi L, Thomson CC, Inoue Y, Johkoh T, et al. Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. Am J Respir Crit Care Med. 2022;205(9):e18-e47 doi.org/10.1164/rccm.202202-0399ST.

18. Wang J, Sun L, Nie Y, Duan S, Zhang T, Wang W, et al. Protein Kinase C δ (PKCδ) Attenuates Bleomycin Induced Pulmonary Fibrosis via Inhibiting NF-κB Signaling Pathway. Front Physiol. 2020;11:367 doi.org/10.3389/fphys.2020.00367.

19. Möller A, Jauch-Speer SL, Gandhi S, Vogl T, Roth J, Fehler O. The roles of toll-like receptor 4, CD33, CD68, CD69, or CD147/EMMPRIN for monocyte activation by the DAMP S100A8/S100A9. Front Immunol. 2023;14:1110185 doi.org/10.3389/fimmu.2023.1110185.

20. Farokhzadian J, Mangolian Shahrbabaki P, Bagheri V. S100A12-CD36 axis: A novel player in the pathogenesis of atherosclerosis? Cytokine. 2019;122:154104 doi.org/https://doi.org/10.1016/j.cyto.2017.07.010.

21. Wang J, Lai X, Yao S, Chen H, Cai J, Luo Y, et al. Nestin promotes pulmonary fibrosis via facilitating recycling of TGF-β receptor I. Eur Respir J. 2022;59(5) doi.org/10.1183/13993003.03721-2020.

22. Liu K, Liu D, Feng Y, Zhang H, Zeng D, Liu Q, et al. Spliceosome-associated protein 130: a novel biomarker for idiopathic pulmonary fibrosis. Ann Transl Med. 2020;8(16):986 doi.org/10.21037/atm-20-4404.

23. Wang F, Li P, Li FS. Integrated Analysis of a Gene Correlation Network Identifies Critical Regulation of Fibrosis by lncRNAs and TFs in Idiopathic Pulmonary Fibrosis. Biomed Res Int. 2020;2020:6537462 doi.org/10.1155/2020/6537462.

24. Butenko S, Satyanarayanan SK, Assi S, Schif-Zuck S, Sher N, Ariel A. Transcriptomic Analysis of Monocyte-Derived Non-Phagocytic Macrophages Favors a Role in Limiting Tissue Repair and Fibrosis. Front Immunol. 2020;11:405 doi.org/10.3389/fimmu.2020.00405.

25. Leurs P, Lindholm B. The AGE-RAGE pathway and its relation to cardiovascular disease in patients with chronic kidney disease. Arch Med Res. 2013;44(8):601-10 doi.org/10.1016/j.arcmed.2013.11.002.

26. Ridley AJ, Schwartz MA, Burridge K, Firtel RA, Ginsberg MH, Borisy G, et al. Cell migration: integrating signals from front to back. Science. 2003;302(5651):1704-9 doi.org/10.1126/science.1092053.

27. Li Y, He Y, Chen S, Wang Q, Yang Y, Shen D, et al. S100A12 as Biomarker of Disease Severity and Prognosis in Patients With Idiopathic Pulmonary Fibrosis. Front Immunol. 2022;13:810338 doi.org/10.3389/fimmu.2022.810338.

28. Simard JC, Cesaro A, Chapeton-Montes J, Tardif M, Antoine F, Girard D, et al. S100A8 and S100A9 induce cytokine expression and regulate the NLRP3 inflammasome via ROS-dependent activation of NF-κB(1.). PLoS One. 2013;8(8):e72138 doi.org/10.1371/journal.pone.0072138.

29. Ryckman C, Vandal K, Rouleau P, Talbot M, Tessier PA. Proinflammatory activities of S100: proteins S100A8, S100A9, and S100A8/A9 induce neutrophil chemotaxis and adhesion. J Immunol. 2003;170(6):3233-42 doi.org/10.4049/jimmunol.170.6.3233.

30. Wilkinson A, Kawaguchi N, Geczy C, Di Girolamo N. S100A8 and S100A9 proteins are expressed by human corneal stromal dendritic cells. Br J Ophthalmol. 2016;100(9):1304-8 doi.org/10.1136/bjophthalmol-2016-308827.

31. von Wulffen M, Luehrmann V, Robeck S, Russo A, Fischer-Riepe L, van den Bosch M, et al. S100A8/A9-alarmin promotes local myeloid-derived suppressor cell activation restricting severe autoimmune arthritis. Cell Rep. 2023;42(8):113006 doi.org/10.1016/j.celrep.2023.113006.

32. Palmer LD, Maloney KN, Boyd KL, Goleniewska AK, Toki S, Maxwell CN, et al. The Innate Immune Protein S100A9 Protects from T-Helper Cell Type 2–mediated Allergic Airway Inflammation. American Journal of Respiratory Cell and Molecular Biology. 2019;61(4):459-68 doi.org/10.1165/rcmb.2018-0217OC.

33. Aggarwal NR, King LS, D'Alessio FR. Diverse macrophage populations mediate acute lung inflammation and resolution. American Journal of Physiology - Lung Cellular and Molecular Physiology.

2014;306(8):L709-L25 doi.org/10.1152/ajplung.00341.2013.

34. Sun SN, Ni SH, Li Y, Liu X, Deng JP, Chen ZX, et al. G-MDSCs promote aging-related cardiac fibrosis by activating myofibroblasts and preventing senescence. Cell Death Dis. 2021;12(6):594 doi.org/10.1038/s41419-021-03874-7.

35. Zhong A, Xu W, Zhao J, Xie P, Jia S, Sun J, et al. S100A8 and S100A9 Are Induced by Decreased Hydration in the Epidermis and Promote Fibroblast Activation and Fibrosis in the Dermis. American Journal of Pathology. 2016;186(1):109-22 doi.org/10.1016/j.ajpath.2015.09.005.

36. Demirdöğen E, Görek Dilektaşlı A, Acet Öztürk NA, Yeşilbursa D, Budak F, Coşkun F, et al. Serum Krebs von den Lungen-6: promising biomarker to differentiate CPFE from IPF. Sarcoidosis Vasc Diffuse Lung Dis. 2022;39(4):e2022035 doi.org/10.36141/svdld.v39i4.11344.

37. Yokoyama A, Kondo K, Nakajima M, Matsushima T, Takahashi T, Nishimura M, et al. Prognostic value of circulating KL-6 in idiopathic pulmonary fibrosis. Respirology. 2006;11(2):164-8 doi.org/10.1111/j.1440-1843.2006.00834.x.

Figures

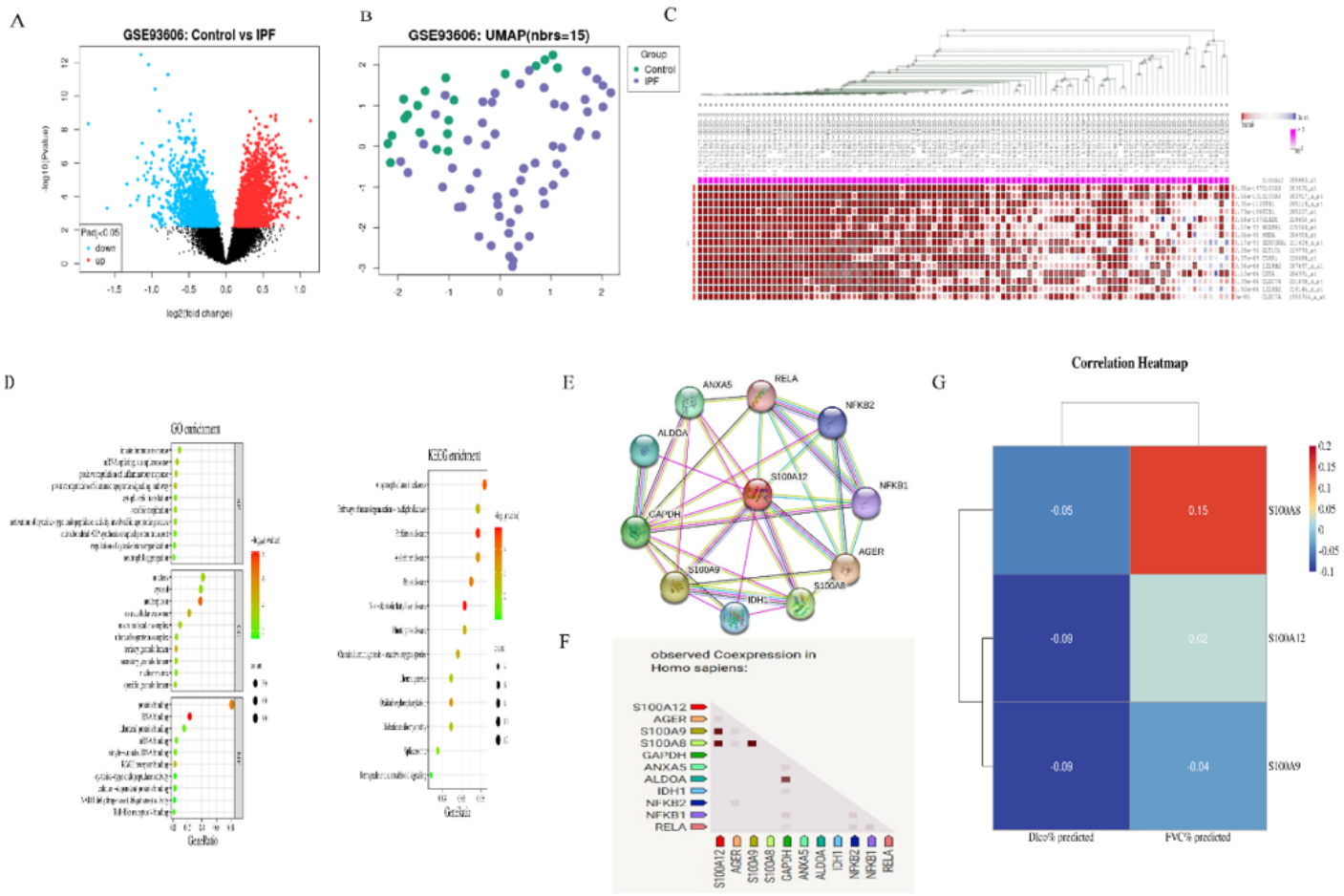

Figure 1

(A) Differentially expressed genes in the IPF group and control group in dataset GSE93606. (B) UMAP cluster of IPF patients. (C) According to the MEM database, the correlation between *S100A12* and 10 genes came from the PPI network. (D) GO and KEGG pathway analysis of *S100A12* and 10 interacting genes. (E) Protein-protein interaction (PPI) network based on STRING database S100A12. (F) Co-expression scores based on RNA expression patterns, and on protein co-regulation provided by Proteome HD. (G) There was no significant association between serum S100A12, S100A8, and S100A9 and lung function.

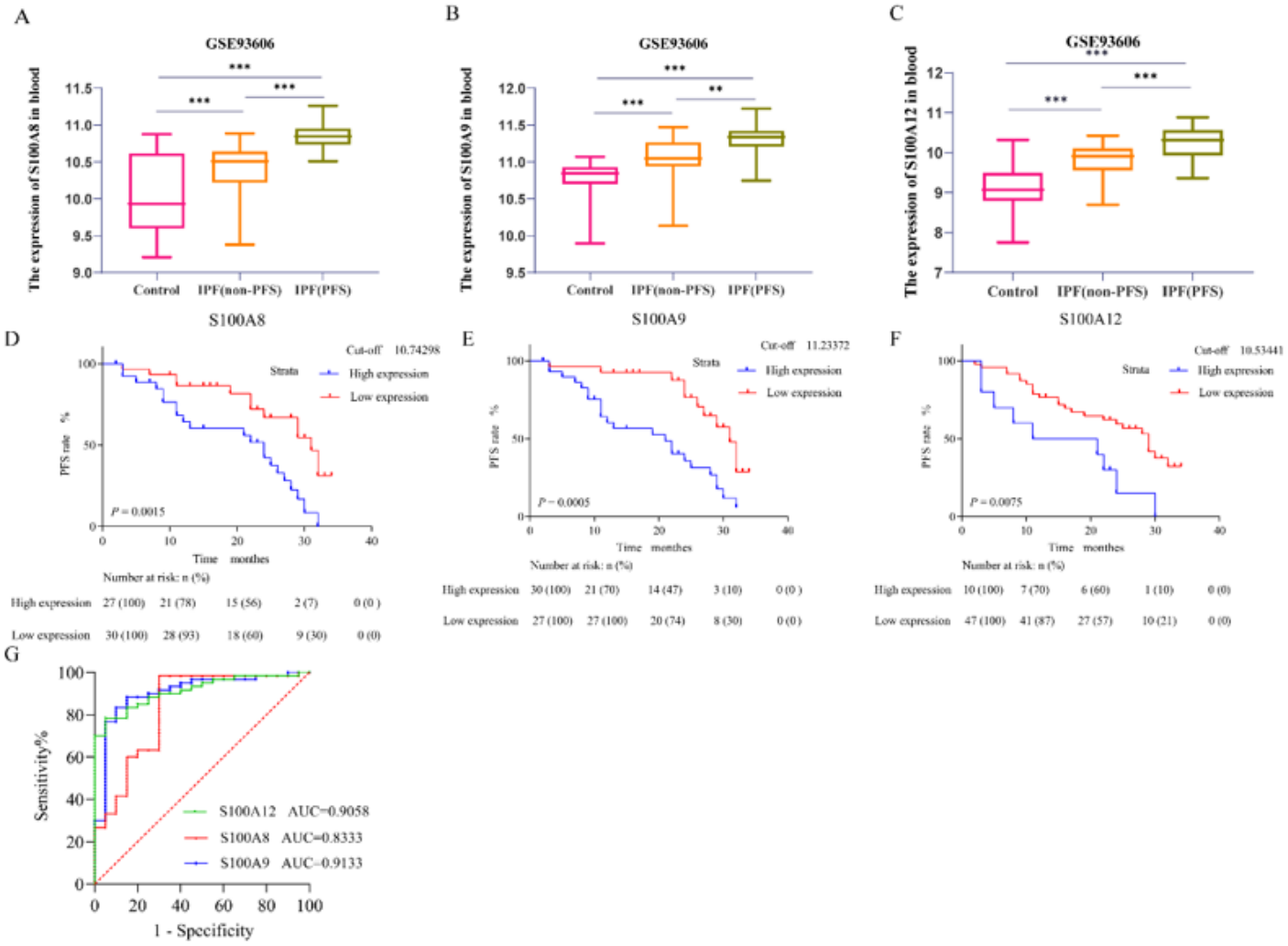

Figure 2

(A, B, C) Expression of three biomarkers (S100A12, S100A8, and S100A9) in the blood of patients with IPF in dataset GSE93606. (D, E, F) Kaplan-Meier survival curve analysis showed that IPF patients with high expression of S100A12, S100A9, and S100A8 in serum had poorer overall survival. (G) Receiver operating characteristic (ROC) curve analysis of peripheral blood S100A8, S100A9, and S100A12 evaluated IPF patients.

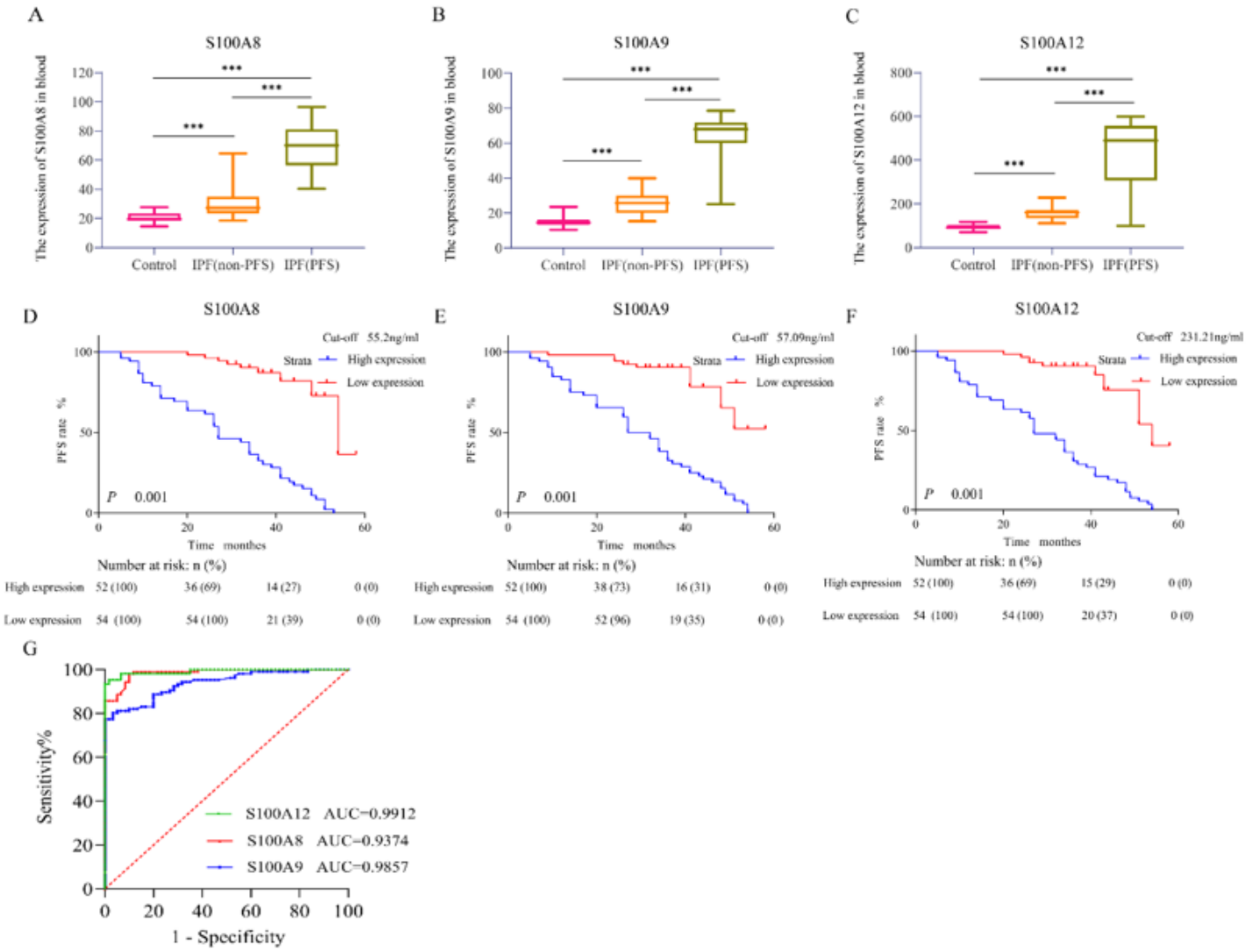

Figure 3

Expression of S100A8 (A), S100A9 (B), and S100A12 (C) in the blood of patients with IPF. (D, E, F) Kaplan-Meier survival curve analysis showed that IPF patients with high peripheral blood expression of S100A12, S100A9, and S100A8 had poor overall survival. (G) ROC curve analysis of peripheral blood S100A8, S100A9, and S100A12 evaluated IPF patients.

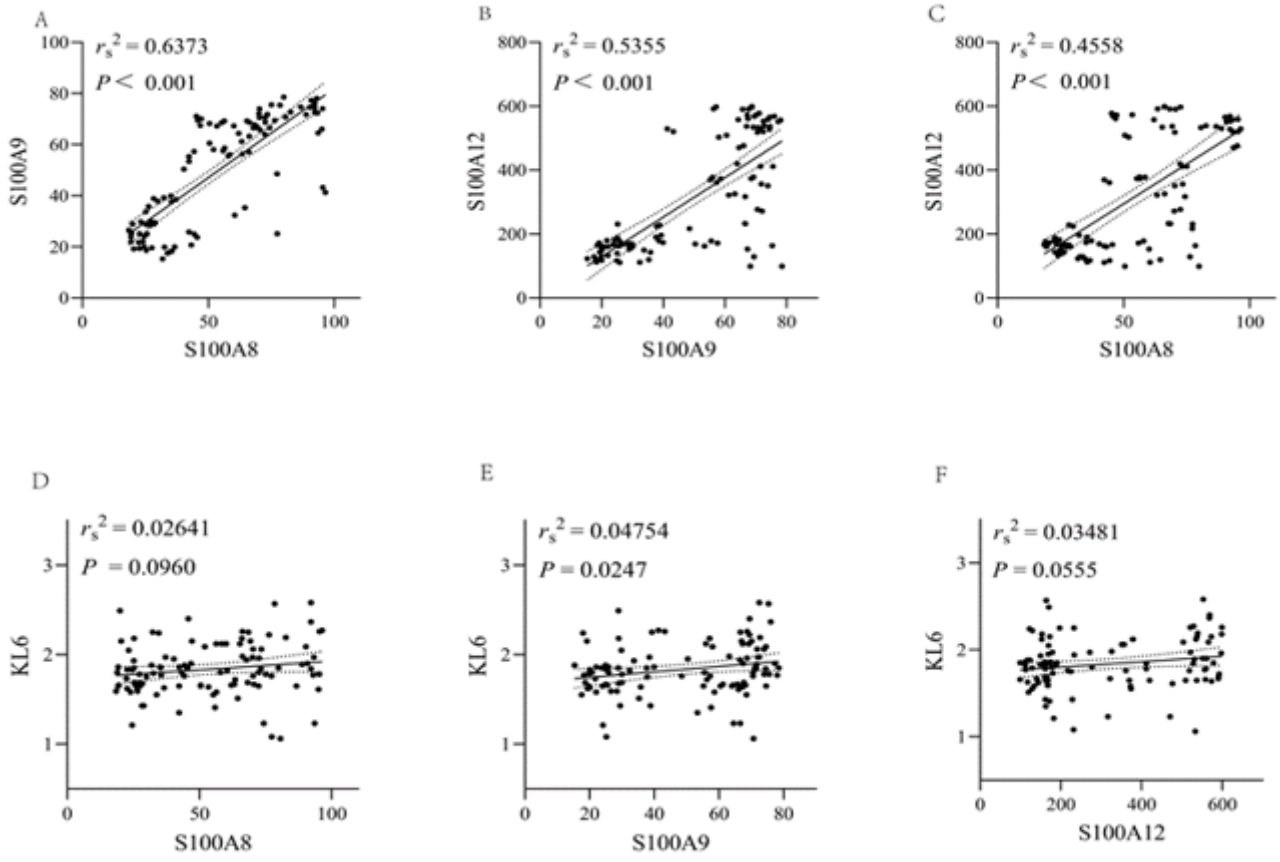

Figure 4

Correlation of serum S100A12 and S100A8/9 levels of IPF with several clinical parameters. There was a strong positive correlation between serum S100A9 and S100A8 levels (Fig. 4A, r_s^2 = 0.6373, P < 0.001). Serum S100A12 and S100A9 levels showed a strong positive correlation (Fig. 4B, r_s^2 = 0.5355, P < 0.001). Strong positive correlation between serum S100A12 and S100A8 levels (Fig. 4C, r_s^2 = 0.4558, P < 0.001). Serum levels of S100A8, S100A9 and S100A12 in IPF patients were positively correlated with KL6 (Fig. 4D, r_s^2 = 0.02641, P = 0.0960; Fig. 4E, r_s^2 = 0.04754, P = 0.0247; Fig. 4F, r_s^2 = 0.03481, P = 0.0555).