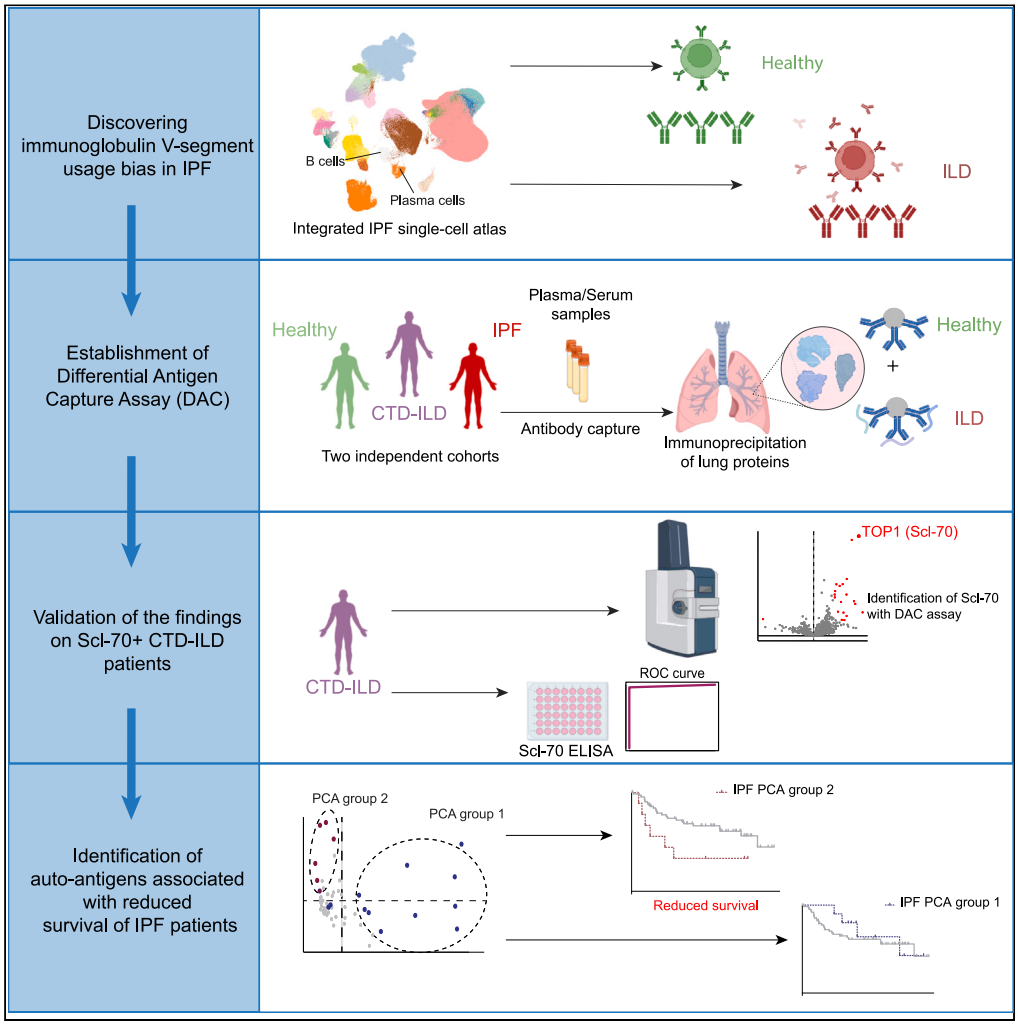
iScience

CellPress
OPEN ACCESS

Article

Mass spectrometry-based autoimmune profiling reveals predictive autoantigens in idiopathic pulmonary fibrosis



Gabriela Leuschner, Anna Semenova, Christoph H. Mayr, ..., Jürgen Behr, Matthias Mann, Herbert B. Schiller

herbert.schiller@helmholtz-munich.de

Highlights

Development of a mass spectrometry-based assay for autoantibody profiling

Multimodal analysis shows biased V-segment usage in IPF

High prevalence of autoreactivities in IPF compared to CTD-ILD

Autoantibodies against THBS1 associated with poor prognosis

Leuschner et al., iScience 26, 108345
November 17, 2023
https://doi.org/10.1016/j.isci.2023.108345

iScience

CellPress
OPEN ACCESS

Article

Mass spectrometry-based autoimmune profiling reveals predictive autoantigens in idiopathic pulmonary fibrosis

Gabriela Leuschner,[1,2,8] Anna Semenova,[1,8] Christoph H. Mayr,[1,8] Theodore S. Kapellos,[1] Meshal Ansari,[1] Benjamin Seeliger,[3] Marion Frankenberger,[1,2] Nikolaus Kneidinger,[2] Rudolf A. Hatz,[4] Anne Hilgendorff,[1,5] Antje Prasse,[3] Jürgen Behr,[2] Matthias Mann,[6] and Herbert B. Schiller[1,7,9,*]

SUMMARY

Autoimmunity plays a role in certain types of lung fibrosis, notably connective tissue disease-associated interstitial lung disease (CTD-ILD). In idiopathic pulmonary fibrosis (IPF), an incurable and fatal lung disease, diagnosis typically requires clinical exclusion of autoimmunity. However, autoantibodies of unknown significance have been detected in IPF patients. We conducted computational analysis of B cell transcriptomes in published transcriptomics datasets and developed a proteomic Differential Antigen Capture (DAC) assay that captures plasma antibodies followed by affinity purification of lung proteins coupled to mass spectrometry. We analyzed antibody capture in two independent cohorts of IPF and CTL-ILD patients over two disease progression time points. Our findings revealed significant upregulation of specific immunoglobulins with V-segment bias in IPF across multiple cohorts. We identified a predictive autoimmune signature linked to reduced transplant-free survival in IPF, persisting over time. Notably, autoantibodies against thrombospondin-1 were associated with decreased survival, suggesting their potential as predictive biomarkers.

INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is currently an incurable disease with a fatal prognosis of 2–5 years after diagnosis.[1] Although newly approved antifibrotic therapies (pirfenidone and nintedanib) can slow down disease progression, the only definitive therapy for IPF is lung transplantation.[2–4] Autoimmunity-driven interstitial lung diseases, such as connective tissue disease-related interstitial lung disease (CTD-ILD), have a better prognosis and are characterized by the presence of auto-antibodies (e.g., Scl-70 antibody against topoisomerase-I in systemic sclerosis), which are used for diagnosis and classification of the patients. While the absence of inflammation and autoimmunity is believed to be characteristic of IPF, there is a growing body of evidence that IPF may have clinical features compatible with an underlying autoimmune process. Circulating autoantibodies have been reported to be present in approximately one-third of IPF patients,[5–7] without meeting the required diagnostic criteria for CTD-ILD and interstitial pneumonia with autoimmune features (IPAF).[8] Given the highly diverse clinical courses of IPF, comprehensive autoantibody profiles from IPF patients may benefit from personalized therapy decisions. Furthermore, it is conceivable that the presence of autoantibodies and autoreactive T cells against unknown antigens could perpetuate pathology in IPF.

Compared to healthy controls and patients with chronic obstructive pulmonary disease (COPD), patients with IPF show higher plasma levels of the B lymphocyte stimulating factor, which is essential for B cell differentiation and B cell survival as well as more C-X-C motif chemokine 13 (CXCL13), mediating B cell recruitment to inflammatory tissue.[9,10] Furthermore, we previously identified an unexpectedly high prevalence of marginal zone B cell-1 protein (MZB1) positive antibody-secreting plasma B cells in ILD tissues, including IPF.[11] In our study, MZB1 levels correlated positively with tissue IgG and negatively with lung function parameters, suggesting a common association of IPF progression

[1]Institute of Lung Health and Immunity, Comprehensive Pneumology Center with the CPC-M bioArchive, Helmholtz Zentrum München, Member of the German Center for Lung Research (DZL), Munich, Germany
[2]Department of Internal Medicine V, Ludwig-Maximilian University Munich, CPC-M bioArchive, Munich, Asklepios Clinics, Gauting, Germany
[3]Department of Pneumology, Hannover Medical School, Member of the German Center for Lung Research (DZL), Hannover, Germany
[4]Center for Thoracic Surgery Munich, Ludwig-Maximilians-University of Munich (LMU), Munich, and Asklepios Medical Center, Member of the German Center for Lung Research (DZL), Gauting, Germany
[5]Center for Comprehensive Developmental Care (CDeCLMU), Hospital of the Ludwig-Maximilians University (LMU), Member of the German Center for Lung Research (DZL), CPC-M bioArchive, Munich, Germany
[6]Max Planck Institute of Biochemistry, Department of Proteomics and Signal Transduction, Martinsried, Germany
[7]Institute of Experimental Pneumology, LMU University Hospital, Ludwig-Maximilians University, Munich, Germany
[8]These authors contributed equally
[9]Lead contact
*Correspondence: herbert.schiller@helmholtz-munich.de
https://doi.org/10.1016/j.isci.2023.108345

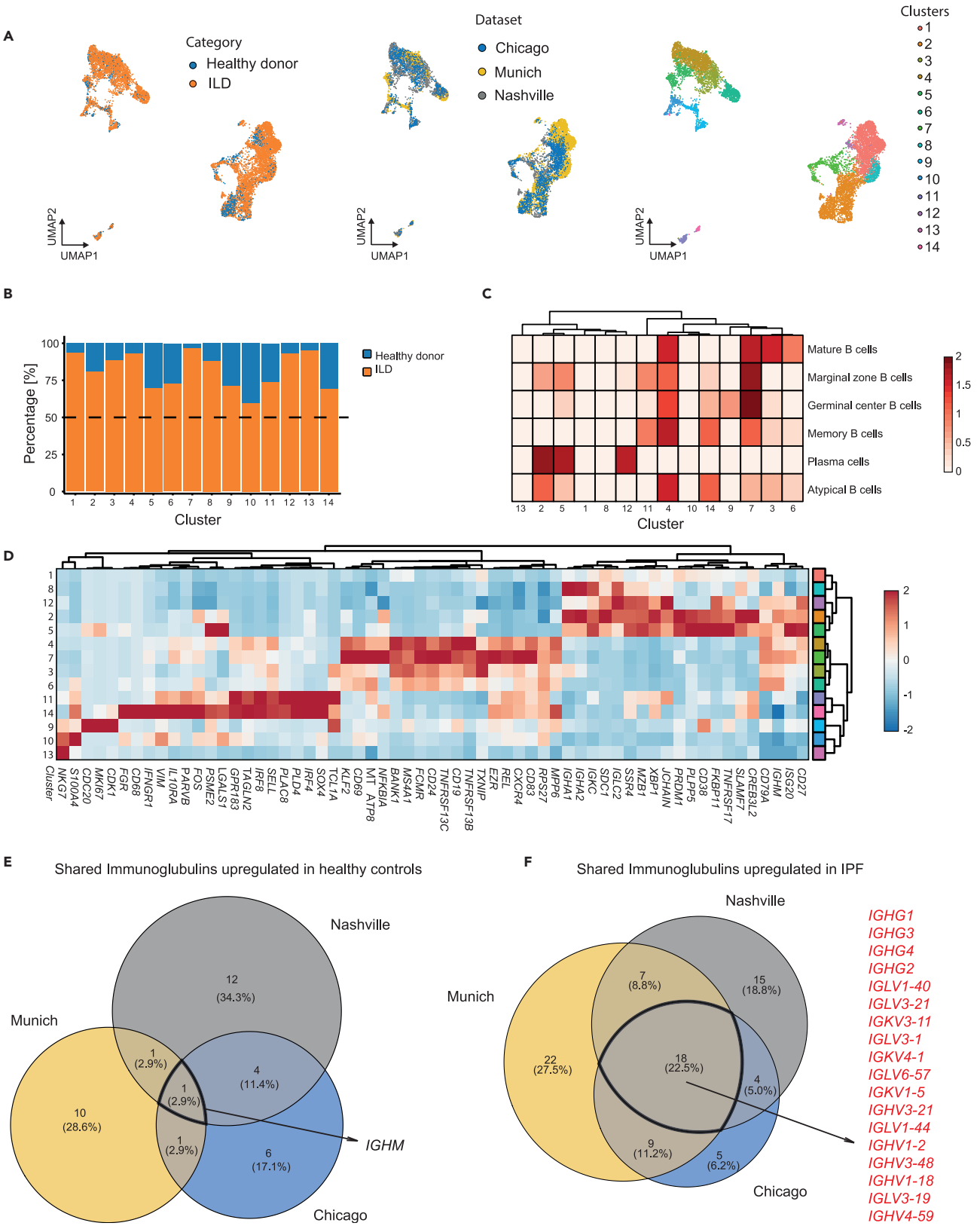

A
Category
Healthy donor
ILD
Dataset
Chicago
Munich
Nashville
Clusters
1
2
3
4
5
6
7
8
9
10
11
12
13
14
UMAP2
UMAP1
B
Percentage [%]
Healthy donor
ILD
Cluster
C
Mature B cells
Marginal zone B cells
Germinal center B cells
Memory B cells
Plasma cells
Atypical B cells
Cluster
D
Cluster
NKG7
S100A4
CDC20
MKI67
CDK1
FGR
CD68
IFNGR1
VIM
IL10RA
PARVB
FOS
PSME2
LGALS1
GPR183
TAGLN2
IRF8
SELL
PLAC8
PLD4
IRF4
SOX4
TCL1A
KLF2
CD69
MT ATP8
NFKBIA
BANK1
MS4A1
FCMR
CD24
TNFRSF13C
CD19
TNFRSF13B
TXNIP
EZR
REL
CXCR4
CD83
RPS27
MPP6
IGHA1
IGHA2
IGKC
SDC1
IGLC2
SSR4
MZB1
XBP1
JCHAIN
PRDM1
PLPP5
CD38
FKBP11
TNFRSF17
SLAMF7
CREB3L2
CD79A
IGHM
ISG20
CD27
E
Shared Immunoglubulins upregulated in healthy controls
Nashville
Munich
Chicago
12 (34.3%)
1 (2.9%)
4 (11.4%)
1 (2.9%)
10 (28.6%)
1 (2.9%)
6 (17.1%)
IGHM
F
Shared Immunoglubulins upregulated in IPF
Nashville
Munich
Chicago
7 (8.8%)
15 (18.8%)
22 (27.5%)
18 (22.5%)
4 (5.0%)
9 (11.2%)
5 (6.2%)
IGHG1
IGHG3
IGHG4
IGHG2
IGLV1-40
IGLV3-21
IGKV3-11
IGLV3-1
IGKV4-1
IGLV6-57
IGKV1-5
IGHV3-21
IGLV1-44
IGHV1-2
IGHV3-48
IGHV1-18
IGLV3-19
IGHV4-59

Figure 1. Pulmonary fibrosis lungs present with elevated B cell numbers and biased V-segment usage

(A) UMAP plots of B/plasma cells mapped by disease group (left), originating dataset (middle) and cluster identity (right).

(B) Bar plot with B cell cluster frequencies in each disease group.

(C) Heatmap of B cell gene signature enrichments in each cluster.

(D) Expression heatmap of B cell gene markers (from C) in each cluster.

(E and F) Venn diagrams showing the overlap of up-regulated most variable immunoglobulin segments between the scRNA-seq datasets of healthy control and pulmonary fibrosis donors.

with antibody-mediated autoimmunity.[11] Indeed, it has been shown that high levels of autoantibodies against vimentin,[12] heat shock protein 70,[13] periplakin as well as anti-parietal cell antibodies,[14,15] were associated with more severe disease in IPF.

In this study, we aimed at the characterization of B cell immunity in lung fibrosis and the identification of disease-specific and prognosis-relevant autoantibodies in IPF and CTD-ILD. We analyzed V-segment usage in antibodies, developed a proteomic workflow for high sensitivity and accuracy detection of lung disease-associated autoantigens, and compared autoreactivity profiles in independent CTD-ILD and IPF patient cohorts longitudinally. Autoreactivities to THBS1 identified a stable patient endotype with reduced survival independently in both study cohorts, suggesting that clinical translation of unbiased autoantibody profiling is highly relevant for personalized IPF patient care.

RESULTS

Immunoglobulin V-segment usage bias in IPF

Immunoreceptors are generated by somatic recombination of multiple gene segments in the genome (VDJ recombination). The selection of V-segments has been shown to be often not random but rather seems to be driven by the nature of the antigen. Thus, V-segment usage may report a prior history of infections/exposures that led to clonal expansion and therefore a biased usage of particular V-segments, as shown for COVID-19 and influenza.[16,17] Importantly, a biased V-segment usage has been also observed previously in various systemic autoimmune diseases, such as systemic lupus erythematosus (SLE).[18]

We were interested in V-segment usage in IPF/ILD and hypothesized that the increased prevalence of plasma B cells in IPF lung tissue[11] could be associated with an increase in autoreactive antibodies. To analyze the B cell repertoire and V-segment usage in ILD across several independent patient cohorts we used publicly available scRNA-seq datasets (Figure S1A) and re-analyzed 568,139 single cells from an integrated ILD cell atlas.[19] Re-clustering of 9,386 B and plasma cells identified 14 clusters that were present in all three datasets (Figures 1A and S1B). All B cell clusters were more abundant in the IPF/ILD samples compared to controls (Figures 1B and S1C).

Calculation of the top five significant gene markers uncovered the underlying heterogeneity within the B/plasma cell lineage (Figure S1D). We used commonly used B cell gene markers[20–25] (Table S1), explored their expression levels in the identified clusters (Figures 1C and 1D), and annotated plasma cells (clusters 1, 2, 5, 8, and 12), mature B cells (clusters 3, 4, 6, and 7), germinal center B cells (cluster 9), and memory B cells (cluster 11). A mature B cell population characterized by the expression of *IFNGR1*, *CD68*, *FGR*, *IL10RA*, *FOS*, and *PARVB*, was reminiscent of the recently described "atypical" B cell state in COVID-19[26] (cluster 14) (Figure 1D).

To analyze potential V-segment usage bias in the IPF/ILD samples in comparison to control donor lungs, we isolated the 10% most variable immunoglobulin genes in each dataset and visualized the agreement between cohorts for genes higher in healthy controls (Figure 1E), or higher in IPF tissues (Figure 1F). While the healthy control donors did not show any conserved V-segment usage across the three patient cohorts, we identified 18 transcripts, including 4 IgG heavy chains (*IGHG1*, *IGHG2*, *IGHG3*, and *IGHG4*) and 14 V-segments (*IGLV1-40*, *IGLV3-21*, *IGKV3-11*, *IGLV3-1*, *IGKV4-1*, *IGLV6-57*, *IGKV1-5*, *IGHV3-21*, *IGLV1-44*, *IGHV1-2*, *IGHV3-48*, *IGHV1-18*, *IGLV3-19*, and *IGHV4-59*), that were reproducibly increased in ILD/IPF across all cohorts.

Taken together, our analysis indicates a significantly biased usage of V-segments in IPF/ILD that either points toward a shared history of specific pathogenic exposures or the presence of a particular repertoire of autoreactive plasma B cell clones.[18]

A mass spectrometry workflow for the proteome-wide discovery of autoantigens in lung disease

We wanted to explore the potential autoreactivity of plasma B cell clones in IPF/ILD using an unbiased quantitative method. Recent advances approach autoantigen discovery in a large scale and high throughput manner using protein microarrays,[27] or phage display libraries, representing a synthetic human peptidome.[28] These assays do, however, depend on synthetic antigens. We instead developed the Differential Antigen Capture assay (DAC; see STAR Methods for detailed description), which is based on immunoprecipitation of human lung proteome extracts with patient immunoglobulins followed by quantitative shotgun proteomics similar to affinity purification (AP)-MS (Figure 2A). To characterize the sensitivity and specificity of our assay we analyzed plasma from patients with CTD-ILD who had undergone Scl-70 autoreactivity testing by ELISA in their clinical routine (SSc: n = 12; undifferentiated CTD: n = 2; RA: n = 1; Figure 2B). Topoisomerase-I (TOP1) is the protein target of the Scl-70 auto-antibodies.

The DAC assay scores the proteins from native lung extracts as putative autoantigens in patients when they are significantly enriched (FDR < 5%) on beads coupled to plasma antibodies from ILD patients versus the beads coupled to plasma antibodies from healthy controls (Figure 2C). In comparison to patients who had negative results for Scl-70 antibodies (n = 7), our assay identified clinically verified Scl-70-positive patients (n = 8) with exceptionally good specificity and sensitivity, and equally good performance as the clinically approved ELISA test (Figures 2D and 2E). In contrast to the ELISA test our method simultaneously identified several other enriched putative autoantigens in these

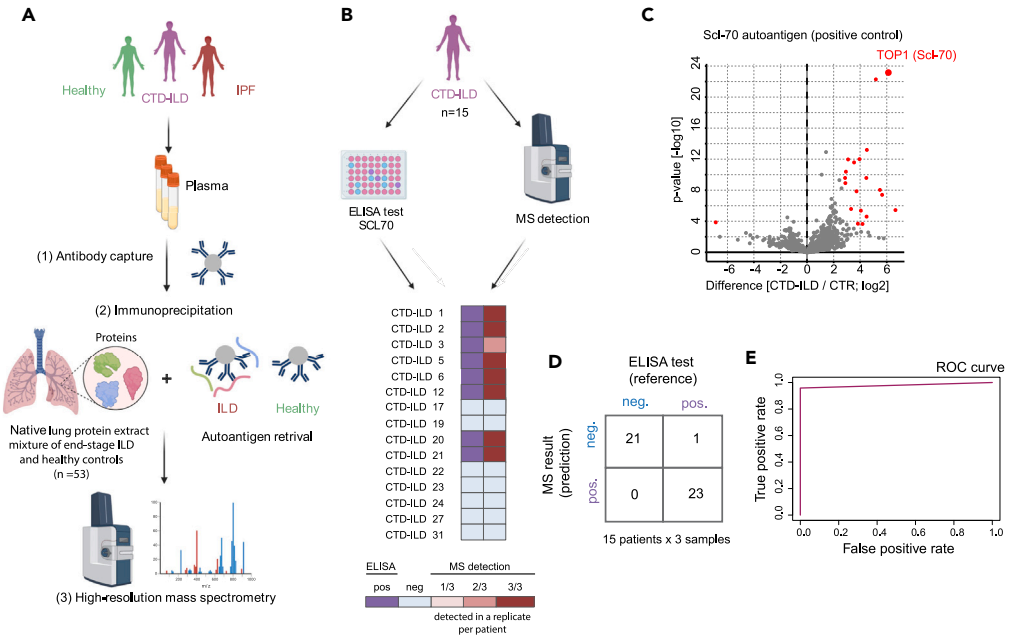


Figure 2. A proteomic workflow detects autoantigens with high sensitivity and specificity

(A) Experimental workflow. Antibodies captured from plasma are used to precipitate proteins from native lung protein extracts. Differential protein binding to beads (disease versus healthy controls) is quantified using shotgun MS-based proteomics.

(B) Detection of plasma antibody reactivity to an autoantigen (SLC-70) by mass spectrometry was compared to a clinically obtained ELISA test in patients with CTD-ILD (SSc: n = 12; undifferentiated CTD: n = 2; RA: n = 1).

(C) Representative volcano plot of one patient with systemic sclerosis (SSc) positive for the Scl-70 antigen (Toposisomerase 1; TOP1) shows enrichment of the autoantigen versus healthy donors. Red dots indicate significantly enriched proteins (FDR < 5%).

(D and E) Each patient sample in panel B was measured in triplicates resulting in the depicted number of true and false positives. The specificity of significant Top1 (Scl-70) enrichment in ILD compared to healthy donor controls by MS analysis was 100%, while sensitivity was 96% (ROC analysis).

patients (Figure 2C). Thus, in this study, we went on to use the multiplexed and unbiased nature of the DAC assay to discover autoantigens in ILD and associate patient-specific signatures with clinical features.

IPF patients feature a high prevalence of autoantibodies

To compare auto-reactivities in CTD-ILD and IPF systematically, we performed the DAC assay in two independent study cohorts. Schematics of cohorts and availability of clinical data are shown in Figure S2. Cohort-1 (Munich; Table S2) included IPF (n = 35), CTD-ILD (n = 26), and healthy controls (n = 31; mean ± SD 58.8 ± 11.1 years, range 40–82 years). The CTD-ILD patients were diagnosed with systemic scleroderma (n = 15), undifferentiated CTD (n = 5), rheumatoid arthritis (RA) (n = 4), Sjogren's syndrome (n = 1) and mixed connective tissue disease (n = 1). The mean ± SD age was 63.3 ± 8.4 years in IPF and 62.7 ± 10.2 years in CTD-ILD. Patients with IPF had significantly lower forced vital capacity (FVC) % pred. than patients with CTD-ILD (54.5 ± 15.8 vs. 65.7 ± 18.9; p = 0.025) (Tables S3 and S4).

The second cohort (Hannover; Table S5) also included IPF (n = 40) and CTD-ILD patients (n = 20; Sjogren's syndrome n = 6; systemic sclerosis n = 5; RA n = 5; CREST syndrome n = 3; antisynthetase syndrome n = 1). On average, with a mean ± SD age of 70.4 ± 8.1 years, IPF patients from cohort-2 were significantly older than all other patients (p = 0.003) and patients from Hannover had significantly better lung function than patients from Munich.

Using the DAC assay we quantified on average >500 proteins per patient sample in all patient groups in both study cohorts (Figure 3A). Putative autoantigens enriched over healthy controls were mostly cytoplasmic, nuclear, and cytoskeletal components (Figures 3B–3D), with infrequent yet biologically relevant reactivity to receptors and secreted proteins (Figure 3E). Overall, 17 antigens were shared across cohorts and disease entities (Figure 3F).

Since the Hannover cohort represented ILD patients at an earlier stage of their disease progression (better lung function), we also performed the DAC assay at a follow-up visit (3.1 years mean time from first sampling with SD of 1.6 years) (Tables S6 and S7). Furthermore, the cohort contained seven healthy serum controls (Table S7). Interestingly, the CTD-ILD and IPF-specific autoantigens were preserved at the second time point with a significant enrichment over control samples, indicating a remarkably stable autoimmune phenotype of these patients (Figures 3G and 3H).

Taking together both cohorts, the mean ± SD number of autoantigens per patient was 16 ± 40 in IPF and 9 ± 15 in CTD-ILD, revealing that surprisingly the prevalence of autoantibodies was not lower in IPF compared to autoimmune-associated ILD (Figure 4A). Independent of

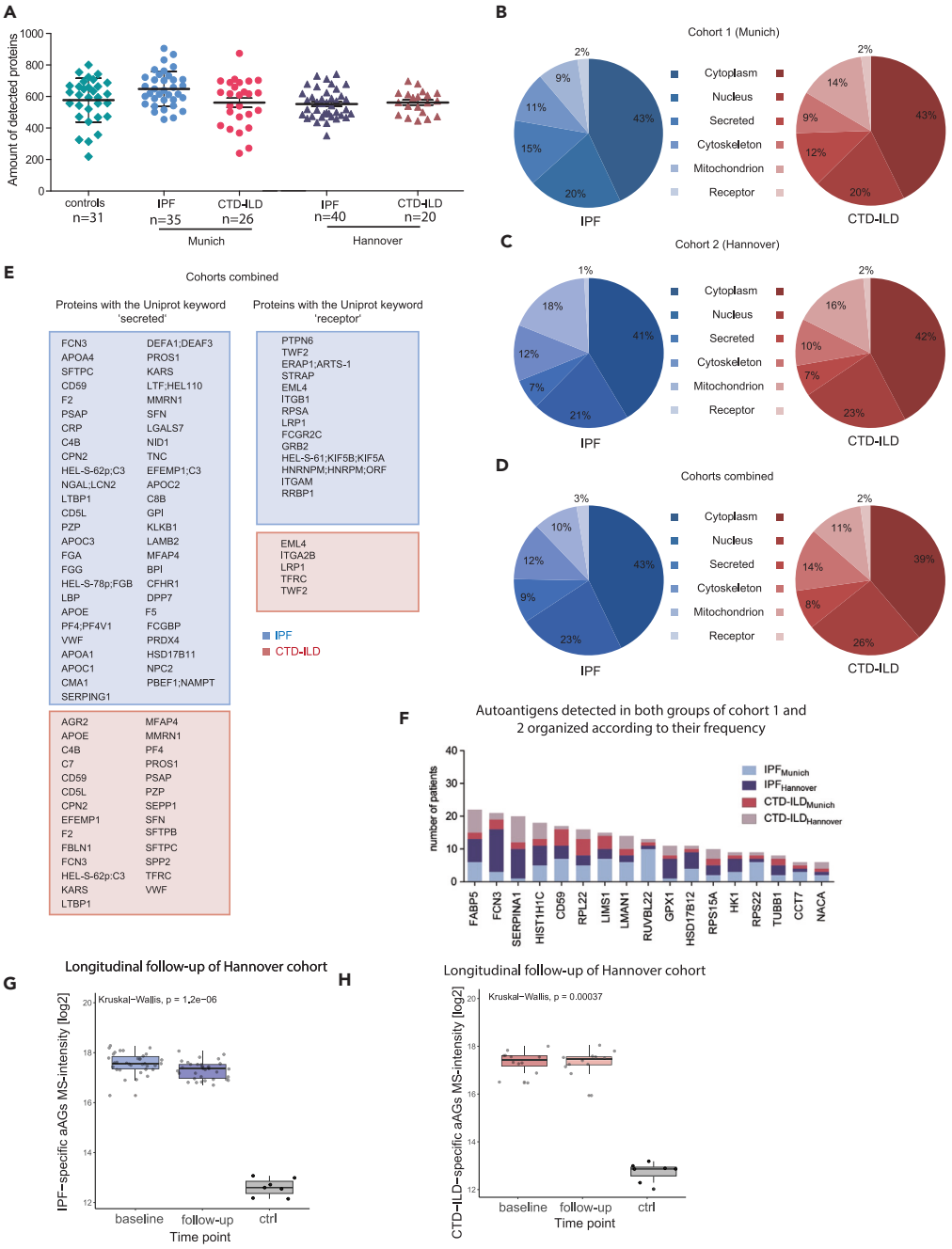


Figure 3. Protein quantification and GO terms of enriched autoantigens

(A) Number of quantified proteins across experimental groups and cohorts.

(B–D) Categories of Autoantigens identified in IPF and CTD-ILD showed similar distribution, in cohort-1 (Munich) (B), cohort-2 (Hannover) (C), and when both cohorts combined (D).

(E) Autoantigens in both cohorts combined, that correspond to the Uniprot keywords "secreted" and "receptor"; blue corresponds to IPF and red to CTL-ILD patients.

(F) Bar graph showing the frequency distribution of 17 antigens found in common in the four groups (IPF and CTD-ILD patients from both cohorts).

(G and H) IPF- and CTD-ILD-specific autoantigens from the table shown in (E) were analyzed at a follow-up sampling as shown (3.1 years mean time from first sampling with SD of 1.6 years), p-value was defined with Kruskal-Wallis test.

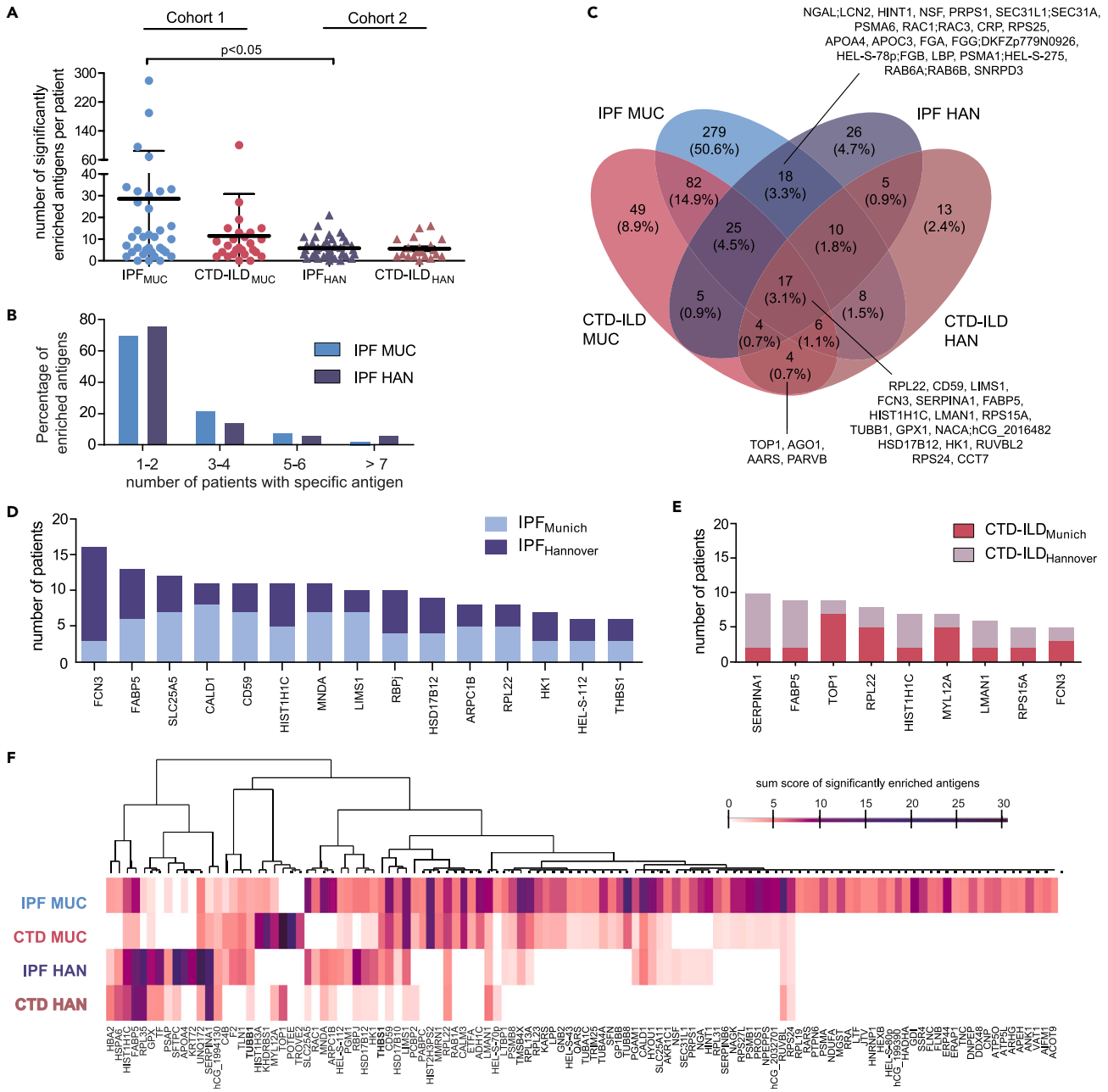


Figure 4. Common and individual autoantigens in IPF and CTD-ILD

(A) Number of significantly enriched autoantigens per patient in cohort-1 and cohort-2.

(B) Percentage distribution of identified autoantigens in IPF cohort-1 and IPF cohort-2.

(C) Venn diagram of individual and shared autoantigens in IPF and CTD-ILD across two cohorts (see Table S8 for Venn diagram loadings).

(D) Bar graph of the 15 most commonly detected autoantigens in IPF (present in at least three patients from each IPF cohort).

(E) Bar graph of the nine most commonly found autoantigens in CTD-ILD (present in at least two patients from each CTD-ILD cohort).

(F) Heatmap of all four groups lists the top identified autoantigens that were significantly enriched in at least three patients in at least one group. Color code shows the score that is the sum of the Student's t test statistics of significant proteins (FDR < 10%).

the cohort, we recognized a broad heterogeneity of autoantigens in IPF: whereas the majority of identified autoreactivities in both IPF cohorts were only found in 1–2 patients, less than 10% of these antigens were enriched in 5 or more patients (Figure 4B).

In IPF patients from cohort-1, the most commonly found autoantigens were RUVBL2, HIST2H3PS2, and TUBB8. While autoreactivity to RUVBL2 and HIST2H3PS2 was not associated with any differences of FVC in IPF patients from cohort-1 (Figures S3A and S3B), patients

with reactivity toward RUVBL2 showed a trend for better survival (p = 0.052; Figure S3D). The detection of TUBB8 autoantibodies was significantly associated with higher FVC and better transplant-free survival in IPF patients from cohort-1 (Figures S3C and S3E). In cohort-2, RUVBL2 and TUBB8 were not frequently detected. In contrast, the most commonly detected autoantigens were FCN3, RPL35 and SERPINA1. None of these three antigens was associated with altered FVC levels (Figure S3F–S3H).

Comparing significantly enriched autoantigens across diagnosis and patient cohorts showed substantial variation with most identified autoantigens specific to either a particular diagnosis or cohort (Figure 4C; Table S8). Some of this variance may be due to cohort specific differences. For instance, the CTD-ILDs from cohort-1 included 62.5% of SSc-ILD patients, while the largest CTD-entity of cohort-2 was Sjogren's syndrome. Nevertheless, a total of 17 autoantigens were shared in all 4 groups (Figure 4C), and 15 autoantigens were present in at least three patients in each IPF cohort, with FCN3 being the most commonly detected (Figure 4D). We further identified 9 autoantigens that were present in at least two patients in each CTD-ILD cohort, including the CTD-ILD specific TOP1 with SERPINA1 having the highest prevalence (Figure 4E). Overall, the autoantigens showed common and distinct associations with study groups (Figure 4F). Of note, IPF patients from cohort-1, which was the clinically most severely affected group, had the largest autoantibody repertoire (Figure 4F). A total of 50.6% (n = 279) of all identified autoantigens were found exclusively in IPF from cohort-1. A high proportion (n = 179; 70.6%) of these individually identified autoantigens were detected in only four patients with IPF from cohort-1 (Figure 4A). These patients appeared to have a distinct autoimmune signature (range 69–279 autoantigens per patient). Clinically, however, there was no evidence of CTD-ILD or IPAF in these patients. In comparison to all other IPF patients from cohort-1, FVC and transplant-free survival were similar in the four IPF patients with distinct autoimmune signature (Figure S4).

We specifically looked out for commonly known autoantigens that have been previously associated with various autoimmune diseases, such as different members of the heat shock protein 70 family (Figure S5). While HSPB1, HSPA5, HSPA12B, and HSPA6 were enriched in IPF, and different types of CTD-ILD, HSPA4, and HSPA9 were only detected in IPF patients. Aminoacyl tRNA synthetase (AARS) was marked as significantly upregulated in one CTD-ILD patient with Antisynthetase syndrome (ASyS) in cohort-2 and additionally in one Systemic sclerosis (SSc) CTD-ILD patient in cohort-1. Regarding Lupus-specific antigens: Ro52 (TRIM21) was significantly upregulated in two SSc patients in cohort-1 and one SSc and one Sjögren's syndrome patient in cohort 2. Additionally, Ro-52 was upregulated in four IPF patients in cohort-1. Further, Ro60 (TROVE2) specific antibodies were upregulated only in CTD-ILD patients in cohort-1: two SSc, one Sjögren's syndrome, and one RA. Autoantibodies against SSB were present in only one RA patient in cohort-1. Finally, auto-reactivities against Fodrin (SPTAN1/SPTBN1) were found in one patient with mixed connective tissue disease in cohort-1; however, they were also significantly upregulated in IPF patients in both cohorts. Anti-centromere antibodies were not detected in any of the cohorts.

In summary, our analysis identified a large number of IPF and CTD-ILD specific and overlapping autoantigens of both intracellular and extracellular nature that persisted in longitudinal analysis. The functional significance of these autoreactivities is currently unclear.

An autoantigen signature is associated with transplant-free survival in IPF

To assess the predictive value of the autoantigens discovered in this study we performed Kaplan-Meier survival analysis for 58 autoantigens present in at least 3 patients. The mean ± SD follow-up in the study cohort was 1.7 ± 1.1 years. While in cohort-1, 4 patients died and 17 underwent lung transplantation, in cohort-2, 10 patients died and none underwent lung transplantation.

Autoreactivities predicted improved and worsened clinical outcomes. Patients with autoreactivities toward APOA4 or GPX1 showed better transplant-free survival (Figures 5A and 5B). Autoreactivity to the ECM glycoprotein thrombospondin-1 (THBS1) predicted a worse survival independently in both study cohorts (Figures 5C and 5D). In a combined analysis of both cohorts, we additionally found THBS1 (n = 6; 8.0%; p = 0.002; Figure 5E), TUBB1 (n = 5; 6.7%; p = 0.019; Figure 5F), and CD5L (n = 3; 4.0%; p = 0.0015; Figure 5G) as being associated with significantly reduced transplant-free survival time. In multiple regression analysis controlling for the confounding age, FVC and gender, both THBS1 and TUBB1 were significant predictors for transplant-free survival (THBS1 HR: 3.98; 95%-CI: 1.432–11.074; p = 0.008; TUBB1, HR 3.469; 95%-CI: 1.150–10.464; p = 0.027).

Patients with autoreactivity for THBS1 often shared autoreactivity with TUBB1, which hinted toward a shared autoantigen repertoire in IPF patients at risk for mortality/lung transplantation. Displaying autoreactivities of IPF patients from both cohorts in a heatmap (Figure 5H), we detected two major clusters. Additionally, PCA analysis of the Z-scored Student's t test statistics of the top 58 autoantigens identified two groups (Figures 5I and 5J). All of the patients in group_1 (n = 11) were from the heatmap cluster-1 and six of nine patients (66.7%) from group_2 were from the heatmap cluster-2.

Based on the PCA groups, we again performed Kaplan-Meier survival curves for transplant-free survival. While patients from PCA-group-1 (cohort-1: n = 8; cohort-2: n = 3) did not show differences in comparison to all other IPF patients (Figure 5K), patients from PCA-group 2 (cohort-1: n = 4; cohort-2: n = 5) had a significantly shorter transplant-free survival in comparison to all other IPF patients from cohort-1 and -2 (median survival 0.9 years versus 3.4 years; p = 0.046; Figure 5L). Of note, these patients did not differ significantly in terms of baseline FVC (p = 0.155) and age (p = 0.282) from the other IPF patients at the time of analysis.

Finally, we measured PCA-group-1 and PCA-group-2 proteins in a longitudinal setting. PCA-group-1 was stably elevated over control samples at the baseline and follow-up time points (Figure 5M, p value < 0.05; follow-up was 3.1 years mean time from first sampling with SD of 1.6 years). PCA-group-2 proteins were also enriched over the control group at both time points with a slight increase observed in the follow-up samples (Figure 5N, p value < 0.05), again indicating that the observed auto activities constitute a remarkably stable and long-lasting phenotype of these patients.

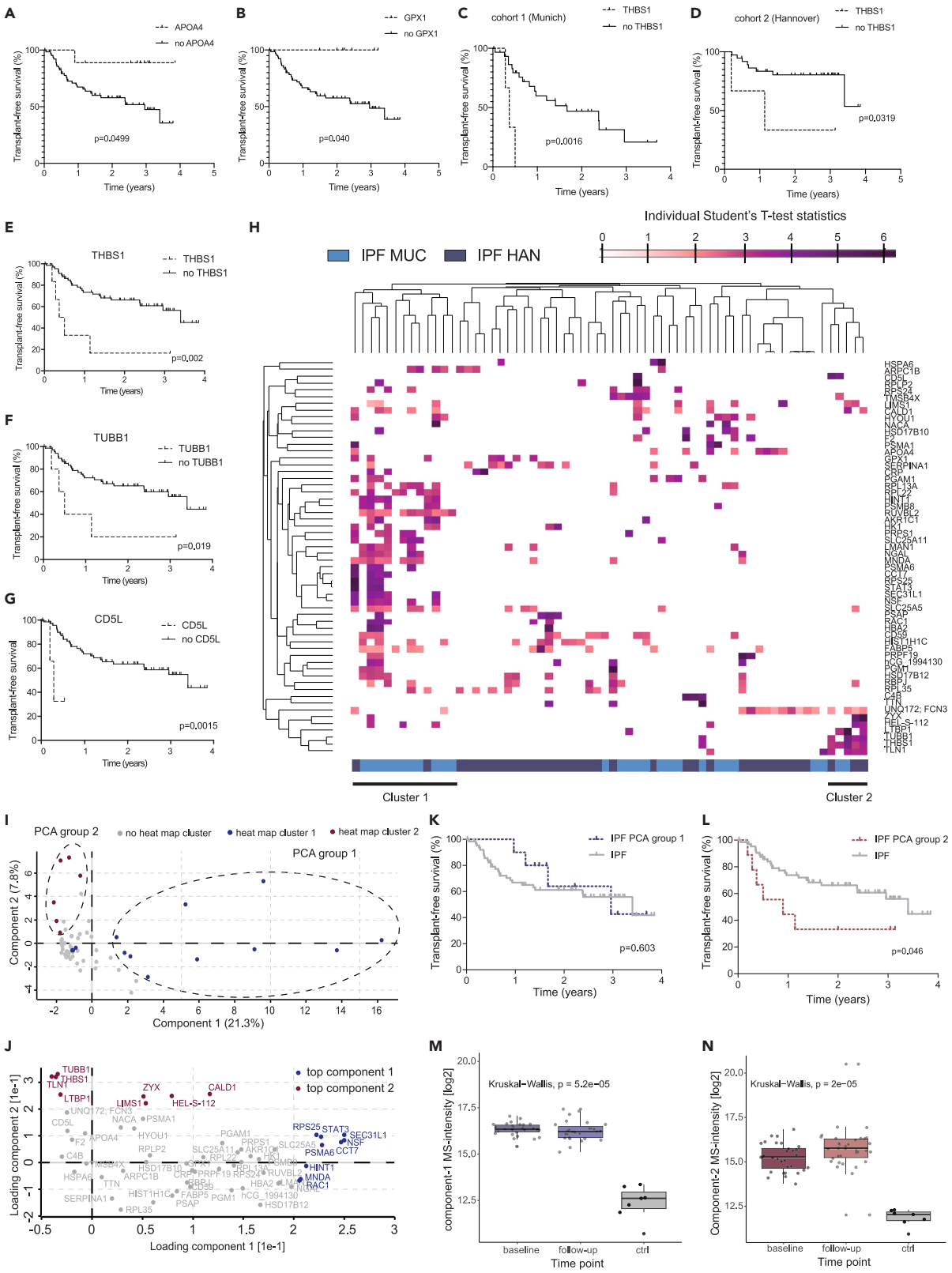

A
APOA4
no APOA4
Transplant-free survival (%)
p=0.0499
Time (years)
B
GPX1
no GPX1
p=0.040
C
cohort 1 (Munich)
THBS1
no THBS1
p=0.0016
D
cohort 2 (Hannover)
p=0.0319
E
THBS1
p=0.002
F
TUBB1
no TUBB1
p=0.019
G
CD5L
no CD5L
p=0.0015
H
IPF MUC
IPF HAN
Individual Student's T-test statistics
Cluster 1
Cluster 2
I
PCA group 2
no heat map cluster
heat map cluster 1
heat map cluster 2
PCA group 1
Component 2 (7.8%)
Component 1 (21.3%)
K
IPF PCA group 1
IPF
p=0.603
L
IPF PCA group 2
p=0.046
J
top component 1
top component 2
Loading component 2 [1e-1]
Loading component 1 [1e-1]
M
Kruskal–Wallis, p = 5.2e−05
component-1 MS-intensity [log2]
baseline
follow-up
ctrl
Time point
N
Kruskal–Wallis, p = 2e−05
Component-2 MS-intensity [log2]

Figure 5. Identification of predictive autoantigens in IPF

(A–G) Kaplan-Meier analysis for transplant-free survival shows a significantly reduced survival in IPF patients with autoantibodies to THBS1, TUBB1, and CD5L respectively.

(H) 58 proteins that were detected in both IPF cohorts and in at least two patients in one of the cohorts were grouped by hierarchical clustering.

(I) Patients were grouped by principal component analysis with the same protein enrichment values used in (H).

(J) Scatterplot showing the loadings of component 1 and component 2 of the PCA in (I).

(K and L) Kaplan-Meier analysis for transplant-free survival does not identify differences in survival in IPF patients from patient group_1 in the PCA (K), but a significantly reduced survival in group_2 (L).

(M and N) PCA group 1 and 2 proteins from the table shown in panel (J) were analyzed at a follow-up sampling as shown (3.1 years mean time from first sampling with SD of 1.6 years), p-value was defined with Kruskal-Wallis test .

DISCUSSION

Developing the mass spectrometry-based DAC assay we were able to globally compare the pulmonary auto-reactivities of plasma antibodies in CTD-ILD and IPF patients, which revealed (1) a surprisingly high prevalence of antibody-mediated autoimmune reactions in IPF compared to CTD-ILD, which coincides with (2) a significant immunoglobulin V-segment bias stable across independent cohorts. Longitudinal analysis showed that (3) patient-specific autoreactivities remained stable over several years, indicating the relevance of these measurements to define patient endotypes at risk. Indeed, we (4) identified autoreactivity profiles, in particular THBS1, that predicted reduced survival in IPF patients, independent of the patient cohort.

Our newly developed autoantigen discovery workflow is highly competitive with current state-of-the-art workflows in several respects. First, we only needed minute amounts of plasma samples (20–50 μL per patient). Second, a non-targeted approach was used. We fed the immunoprecipitation assay with a full representation of the diseased lung proteome and identified antigens using mass spectrometry. Targeted assays depend on recombinant purified proteins, making comprehensive and high throughput screening difficult. Third, autoantigens can carry disease-associated posttranslational protein modifications (PTMs), which also cannot be easily addressed with methods based on recombinant proteins. We are therefore confident that the DAC assay offers several advantages for both basic research and future clinical utility. For instance, the DAC assay can be used to shed light on the currently ongoing discussion on the potential role of autoantibodies in (long) COVID-19.[29]

Autoantibodies against THBS1, which were found in 8% of IPF patients, were an independent predictor for transplant-free survival in IPF patients in both study and validation cohorts when adjusting for baseline age, FVC and gender. THBS1 is an ECM glycoprotein with multiple functions, including the inhibition and stimulation of angiogenesis and tumor progression or wound healing.[30] Interestingly, loss of THBS1 was associated with impaired lung regeneration,[31] suggesting that autoantibodies in patients may cause loss of function phenotypes that promote fibrosis as a consequence of failed alveolar epithelial regeneration. In both study cohorts, we found several patients who partially shared a similar set of autoantigens including THBS1 (PCA-group-2 in Figure 5I) that had a significantly worsened prognosis. This signature of autoantigens is enriched for proteins functional in either the regulation of TGF-beta[32,33] (THBS1 and LTBP1) or fibroblast focal adhesions[34,35] (TLN1, ZYX, and LIMS1) as well as the cytoskeleton (TUBB1 and CALD1). Together, these data suggest that the autoreactivities in PCA-group-2 that were associated with reduced patient survival may modulate the TGF-beta pathway as well as cell shape and adhesion mechanisms.

Our study also identified autoreactivities that were associated with better outcomes (e.g., APOA4 and GPX1). The concept of protective autoimmunity has recently been suggested in the COVID-19 context based on the observation that autoantibodies toward chemokines were associated with favorable disease outcomes and negatively correlated with the development of long COVID at 1 year post-infection.[36,37] Application of the DAC assay to the (long-) COVID-19 context and autoimmunity post-viral infection in general will thus be a very interesting and promising future direction.

In conclusion, the surprisingly high prevalence of autoantibodies in IPF warrants further research on their causal roles in disease progression or their function in protecting the patient. Based on autoimmune reactivity profiles, we identified a molecular endotype of IPF patients at risk that is worth further clinical investigation.

Limitations of the study

We acknowledge the limitations of our study. We made use of two retrospective cohorts that were highly distinct and therefore characterized by important differences. First, IPF patients from cohort-2 were not only older but were less restricted in terms of lung function compared to patients from cohort-1. Moreover, the body fluids we used from cohort-1 were plasma samples and the ones from cohort-2 serum, which might account for some of the heterogeneity between cohorts. The sensitivity of the DAC assay depends on several parameters that can vary for the respective autoantigen/antibody pairs. The enrichment of autoantigen depends on (1) autoantibody titer in plasma, (2) antibody affinity, (3) concentration of autoantigen in the tissue extract, and (4) other unknown criteria. Since we had to choose one defined condition for our experiments we expected to miss some of the possible targets.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.108345.

ACKNOWLEDGMENTS

We thank all the patients and their families for supporting the progress of science. We gratefully acknowledge the provision of human biomaterial and clinical data from the CPC-M bioArchive and its partners at the Asklepios Biobank Gauting, the Klinikum der Universität München and the Ludwig-Maximilians-Universität München. We thank Ina Koch, Britta Peschel, Annika Frank, Anja Disovic, and Marion Frankenberger for managing the Asklepios Biobank and the CPC-M Bioarchive and supporting this study. We thank Igor Paron and Christian Deiml for their expert support of the proteomics pipeline. Also, we thank Lukas Heumos for providing access to the Biorender account. This work was supported by the German Center for Lung Research (DZL), the Helmholtz Association (CoViPa - lessons to get prepared for future pandemics), the Munich Medical & Clinician Scientist Program (MCSP) of the medical faculty of the Ludwig-Maximilians-Universität München and the Max Planck Society.

AUTHOR CONTRIBUTIONS

H.B.S. conceived the experimental design and research narrative and supervised the entire study. H.B.S., G.L., and T.S.K. wrote the paper. C.H.M., A.S., and G.L. performed proteomics experiments. G.L. analyzed clinical data. G.L., C.H.M., A.S., and M.A. analyzed proteomics data. B.S., N.K., M.F., and R.A.H. collected clinical samples. T.S.K. analyzed the single-cell transcriptomics dataset. A.H., A.P., J.B., M.M., and H.B.S. provided resources. All authors read and approved the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 5, 2023
Revised: September 13, 2023
Accepted: October 23, 2023
Published: October 27, 2023

REFERENCES

1. Bjoraker, J.A., Ryu, J.H., Edwin, M.K., Myers, J.L., Tazelaar, H.D., Schroeder, D.R., and Offord, K.P. (1998). Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. Am. J. Respir. Crit. Care Med. *157*, 199–203.
2. King, T.E., Bradford, W.Z., Castro-Bernardini, S., Fagan, E.A., Glaspole, I., Glassberg, M.K., Gorina, E., Hopkins, P.M., Kardatzke, D., Lancaster, L., et al. (2014). A Phase 3 Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. N. Engl. J. Med. *370*, 2083–2092. https://doi.org/10.1056/nejmoa1402582.
3. Noble, P.W., Albera, C., Bradford, W.Z., Costabel, U., Glassberg, M.K., Kardatzke, D., King, T.E., Lancaster, L., Sahn, S.A., Szwarcberg, J., et al. (2011). Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. Lancet *377*, 1760–1769. https://doi.org/10.1016/s0140-6736(11)60405-4.
4. Richeldi, L., du Bois, R.M., Raghu, G., Azuma, A., Brown, K.K., Costabel, U., Cottin, V., Flaherty, K.R., Hansell, D.M., Inoue, Y., et al. (2014). Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. N. Engl. J. Med. *370*, 2071–2082. https://doi.org/10.1056/nejmoa1402584.
5. Lee, J.S., Kim, E.J., Lynch, K.L., Elicker, B., Ryerson, C.J., Katsumoto, T.R., Shum, A.K., Wolters, P.J., Cerri, S., Richeldi, L., et al. (2013). Prevalence and clinical significance of circulating autoantibodies in idiopathic pulmonary fibrosis. Respir. Med. *107*, 249–255.
6. Leuschner, G., Reiter, F., Stocker, F., Crispin, A., Kneidinger, N., Veit, T., Klenner, F., Ceelen, F., Zimmermann, G., Leuchte, H., et al. (2018). Idiopathic Pulmonary Fibrosis Among Young Patients: Challenges in Diagnosis and Management. Lung *196*, 401–408.

7. Moua, T., Maldonado, F., Decker, P.A., Daniels, C.E., and Ryu, J.H. (2014). Frequency and Implication of Autoimmune Serologies in Idiopathic Pulmonary Fibrosis. Mayo Clin. Proc. *89*, 319–326. https://doi.org/10.1016/j.mayocp.2013.11.018.
8. Fischer, A., Antoniou, K.M., Brown, K.K., Cadranel, J., Corte, T.J., du Bois, R.M., Lee, J.S., Leslie, K.O., Lynch, D.A., Matteson, E.L., et al. (2015). An official European Respiratory Society/American Thoracic Society research statement: interstitial pneumonia with autoimmune features. Eur. Respir. J. *46*, 976–987. https://doi.org/10.1183/13993003.00150-2015.
9. Xue, J., Kass, D.J., Bon, J., Vuga, L., Tan, J., Csizmadia, E., Otterbein, L., Soejima, M., Levesque, M.C., Gibson, K.F., et al. (2013). Plasma B Lymphocyte Stimulator and B Cell Differentiation in Idiopathic Pulmonary Fibrosis Patients. J. Immunol. *191*, 2089–2095. https://doi.org/10.4049/jimmunol.1203476.
10. Vuga, L.J., Tedrow, J.R., Pandit, K.V., Tan, J., Kass, D.J., Xue, J., Chandra, D., Leader, J.K., Gibson, K.F., Kaminski, N., et al. (2014). C-X-C motif chemokine 13 (CXCL13) is a prognostic biomarker of idiopathic pulmonary fibrosis. Am. J. Respir. Crit. Care Med. *189*, 966–974.
11. Schiller, H.B., Mayr, C.H., Leuschner, G., Strunz, M., Staab-Weijnitz, C., Preisendörfer, S., Eckes, B., Moinzadeh, P., Krieg, T., Schwartz, D.A., et al. (2017). Deep Proteome Profiling Reveals Common Prevalence of MZB1-Positive Plasma B Cells in Human Lung and Skin Fibrosis. Am. J. Respir. Crit. Care Med. *196*, 1298–1310.
12. Li, F.J., Surolia, R., Li, H., Wang, Z., Kulkarni, T., Liu, G., de Andrade, J.A., Kass, D.J., Thannickal, V.J., Duncan, S.R., and Antony, V.B. (2017). Autoimmunity to Vimentin Is Associated with Outcomes of Patients with Idiopathic Pulmonary Fibrosis. J. Immunol. *199*, 1596–1605.
13. Kahloon, R.A., Xue, J., Bhargava, A., Csizmadia, E., Otterbein, L., Kass, D.J., Bon, J., Soejima, M., Levesque, M.C., Lindell, K.O., et al. (2013). Patients with idiopathic pulmonary fibrosis with antibodies to heat shock protein 70 have poor prognoses. Am. J. Respir. Crit. Care Med. *187*, 768–775.
14. Beltramo, G., Thabut, G., Peron, N., Nicaise, P., Cazes, A., Debray, M.-P., Joannes, A., Castier, Y., Mailleux, A.A., Frija, J., et al. (2018). Anti-parietal cell autoimmunity is associated with an accelerated decline of lung function in IPF patients. Respir. Med. *135*, 15–21.
15. Taillé, C., Grootenboer-Mignot, S., Boursier, C., Michel, L., Debray, M.-P., Fagart, J., Barrientos, L., Mailleux, A., Cigna, N., Tubach, F., et al. (2011). Identification of Periplakin as a New Target for Autoreactivity in Idiopathic Pulmonary Fibrosis. Am. J. Respir. Crit. Care Med. *183*, 759–766. https://doi.org/10.1164/rccm.201001-0076oc.
16. Unterman, A., Sumida, T.S., Nouri, N., Yan, X., Zhao, A.Y., Gasque, V., Schupp, J.C., Asashima, H., Liu, Y., Cosme, C., et al. (2022). Single-cell multi-omics reveals dyssynchrony of the innate and adaptive immune system in progressive COVID-19. Nat. Commun. *13*, 1–23.
17. Avnir, Y., Watson, C.T., Glanville, J., Peterson, E.C., Tallarico, A.S., Bennett, A.S., Qin, K., Fu, Y., Huang, C.-Y., Beigel, J.H., et al. (2016). IGHV1-69 polymorphism modulates anti-influenza antibody repertoires, correlates with IGHV utilization shifts and varies by ethnicity. Sci. Rep. *6*, 20842.
18. Foreman, A.L., Van de Water, J., Gougeon, M.-L., and Gershwin, M.E. (2007). B cells in autoimmune diseases: insights from analyses of immunoglobulin variable (Ig V) gene usage. Autoimmun. Rev. *6*, 387–401.
19. Mayr, C.H., Simon, L.M., Leuschner, G., Ansari, M., Schniering, J., Geyer, P.E., Angelidis, I., Strunz, M., Singh, P., Kneidinger, N., et al. (2021). Integrative analysis of cell state changes in lung fibrosis with peripheral protein biomarkers. EMBO Mol. Med. *13*, e12871.
20. Morgan, D., and Tergaonkar, V. (2022). Unraveling B cell trajectories at single cell resolution. Trends Immunol. *43*, 210–229.
21. He, S., Wang, L.-H., Liu, Y., Li, Y.-Q., Chen, H.-T., Xu, J.-H., Peng, W., Lin, G.-W., Wei, P.-P., Li, B., et al. (2020). Single-cell transcriptome profiling of an adult human cell atlas of 15 major organs. Genome Biol. *21*, 294.
22. King, H.W., Orban, N., Riches, J.C., Clear, A.J., Warnes, G., Teichmann, S.A., and James, L.K. (2021). Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. Sci. Immunol. *6*, eabe6291. https://doi.org/10.1126/sciimmunol.abe6291.
23. Holmes, A.B., Corinaldesi, C., Shen, Q., Kumar, R., Compagno, N., Wang, Z., Nitzan, M., Grunstein, E., Pasqualucci, L., Dalla-Favera, R., and Basso, K. (2020). Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. J. Exp. Med. *217*, e20200483. https://doi.org/10.1084/jem.20200483.
24. Zhang, C., Zhang, T.-X., Liu, Y., Jia, D., Zeng, P., Du, C., Yuan, M., Liu, Q., Wang, Y., and Shi, F.-D. (2021). B-Cell Compartmental Features and Molecular Basis for Therapy in Autoimmune Disease. Neurol. Neuroimmunol. Neuroinflamm. *8*, e1070. https://doi.org/10.1212/NXI.0000000000001070.
25. Stewart, A., Ng, J.C.-F., Wallis, G., Tsioligka, V., Fraternali, F., and Dunn-Walters, D.K. (2021). Single-Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete Development Pathways. Front. Immunol. *12*, 602539.
26. Su, Y., Yuan, D., Chen, D.G., Ng, R.H., Wang, K., Choi, J., Li, S., Hong, S., Zhang, R., Xie, J., et al. (2022). Multiple early factors anticipate post-acute COVID-19 sequelae. Cell *185*, 881–895.e20.
27. Rosenberg, J.M., and Utz, P.J. (2015). Protein Microarrays: A New Tool for the Study of Autoantibodies in Immunodeficiency. Front. Immunol. *6*, 138. https://doi.org/10.3389/fimmu.2015.00138.
28. Larman, H.B., Zhao, Z., Laserson, U., Li, M.Z., Ciccia, A., Gakidis, M.A.M., Church, G.M., Kesari, S., Leproust, E.M., Solimini, N.L., and Elledge, S.J. (2011). Autoantigen discovery with a synthetic human peptidome. Nat. Biotechnol. *29*, 535–541. https://doi.org/10.1038/nbt.1856.
29. Khamsi, R. (2021). Rogue antibodies could be driving severe COVID-19. Nature *590*, 29–31.
30. Bornstein, P. (1995). Diversity of function is inherent in matricellular proteins: an appraisal of thrombospondin 1. J. Cell Biol. *130*, 503–506.
31. Lee, J.H., Bhang, D.H., Beede, A., Huang, T.L., Stripp, B.R., Bloch, K.D., Wagers, A.J., Tseng, Y.H., Ryeom, S., and Kim, C.F. (2014). Lung stem cell differentiation in mice directed by endothelial cells via a BMP4-NFATc1-thrombospondin-1 axis. Cell *156*, 440–455. https://doi.org/10.1016/j.cell.2013.12.039.
32. Murphy-Ullrich, J.E., and Suto, M.J. (2018). Thrombospondin-1 regulation of latent TGF-β activation: A therapeutic target for fibrotic disease. Matrix Biol. *68–69*, 28–43. https://doi.org/10.1016/j.matbio.2017.12.009.
33. Rifkin, D., Sachan, N., Singh, K., Sauber, E., Tellides, G., and Ramirez, F. (2022). The role of LTBPs in TGF beta signaling. Dev. Dyn. *251*, 95–104.
34. Schiller, H.B., Friedel, C.C., Boulegue, C., and Fässler, R. (2011). Quantitative proteomics of the integrin adhesome show a myosin II-dependent recruitment of LIM domain proteins. EMBO Rep. *12*, 259–266.
35. Schiller, H.B., and Fässler, R. (2013). Mechanosensitivity and compositional dynamics of cell-matrix adhesions. EMBO Rep. *14*, 509–519.
36. Muri, J., Cecchinato, V., Cavalli, A., Shanbhag, A.A., Matkovic, M., Biggiogero, M., Maida, P.A., Moritz, J., Toscano, C., Ghovehoud, E., et al. (2023). Autoantibodies against chemokines post-SARS-CoV-2 infection correlate with disease course. Nat. Immunol. *24*, 604–611. https://doi.org/10.1038/s41590-023-01445-w.
37. Wang, E.Y., Mao, T., Klein, J., Dai, Y., Huck, J.D., Jaycox, J.R., Liu, F., Zhou, T., Israelow, B., Wong, P., et al. (2021). Diverse functional autoantibodies in patients with COVID-19. Nature *595*, 283–288. https://doi.org/10.1038/s41586-021-03631-y.
38. Habermann, A.C., Gutierrez, A.J., Bui, L.T., Yahn, S.L., Winters, N.I., Calvi, C.L., Peter, L., Chung, M.-I., Taylor, C.J., Jetter, C., et al. (2020). Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. Sci. Adv. *6*, eaba1972.
39. Reyfman, P.A., Walter, J.M., Joshi, N., Anekalla, K.R., McQuattie-Pimentel, A.C., Chiu, S., Fernandez, R., Akbarpour, M., Chen, C.-I., Ren, Z., et al. (2019). Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. Am. J. Respir. Crit. Care Med. *199*, 1517–1536.
40. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. *47*, D442–D450.
41. Schiller, H.B., Fernandez, I.E., Burgstaller, G., Schaab, C., Scheltema, R.A., Schwarzmayr, T., Strom, T.M., Eickelberg, O., and Mann, M. (2015). Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. Mol. Syst. Biol. *11*, 819.
42. Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat. Methods *11*, 319–324.
43. Thakur, S.S., Geiger, T., Chatterjee, B., Bandilla, P., Fröhlich, F., Cox, J., and Mann, M. (2011). Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. Mol. Cell. Proteomics *10*. M110.003699.
44. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies

and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

45. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of proteomics data. Nat. Methods *13*, 731–740.
46. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J.V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. J. Proteome Res. *10*, 1794–1805.
47. Meier, F., Brunner, A.-D., Koch, S., Koch, H., Lubeck, M., Krause, M., Goedecke, N., Decker, J., Kosinski, T., Park, M.A., et al. (2018). Online Parallel Accumulation–Serial Fragmentation (PASEF) with a Novel Trapped Ion Mobility Mass Spectrometer. Mol. Cell. Proteomics *17*, 2534–2545.
48. Mayr, C.H., Simon, L.M., Leuschner, G., Ansari, M., Schniering, J., Geyer, P.E., Angelidis, I., Strunz, M., Singh, P., Kneidinger, N., et al. (2021). Integrative analysis of cell state changes in lung fibrosis with peripheral protein biomarkers. EMBO Mol. Med. *13*, e12871.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
PhosSTOP	Roche	Cat#4906845001
cOmplete™, EDTA-freier Proteasehemmer-Cocktail	Roche	Cat#11873580001
Lys-C Endoproteinase, MS Grade	Thermo Scientific	Cat# 90051
Trypsin Protease, MS Grade	Thermo Scientific	Cat# 90059
Critical commercial assays		
Pierce BCA Protein Assay Kits	Thermo Scientific	Cat#23225
Deposited data		
DAC assay Munich, cohort 1	This paper	PXD024113
DAC assay Hannover, cohort 2	This paper	PXD024123
Sc-transcriptomic data, IPF and control lung tissue	Mayer et al., 2021	NA
Sc-transcriptomic data, IPF and control lung tissue	Habermann et al.[38]	GSE135893
Sc-transcriptomic data, IPF and control lung tissue	Reyman et al.[39]	https://nupulmonary.org/resources/
Software and algorithms		
MaxQuant software	https://www.maxquant.org/	v1.6.7.0
Perseus software	https://maxquant.net/perseus/	v 1.5.3.0, 1.6.2.3, 1.6.10.50
DEP package	https://bioconductor.org/packages/release/bioc/html/DEP.html	v1.21.0
Seurat	https://satijalab.org/seurat/	v4.1.0
R	https://www.r-project.org/	v4.0.5
Harmony algorithm	https://github.com/immunogenomics/harmony	v0.1.0
AUCell package	https://www.bioconductor.org/packages/release/bioc/html/AUCell.html	v1.12.0
GraphPad Prism	https://www.graphpad.com/features	v5.0
Other		
Bioruptor® Plus sonication device	Digenode	Cat#B01020001
Pierce™ Protein G Agarose	Thermo Scientific	Cat# 20397
TissueLyser 2	Qiagen	NA
MultiScreenHTS BV Filter Plate, 1.2 µm, clear, sterile	Merck	Cat# MSBVS1210
CDS Empore™ SDB-RPS Extraction Disks	Fisher Scientific	Cat# 13-110-022
EASY-nLC™ 1200 System	Thermo Scientific	Cat# LC140
timsTOF Pro	Bruker Daltonik	NA
ReproSil-Pur C18-AQ 1.9 µm resin	Dr.Maisch GmbH	NA
UniProt FASTA database	https://www.uniprot.org/	UP000005640_9606
EASY-nLC 1000 System	Thermo Scientific	Cat# LC120
Q-Exactive Mass Spectrometer	Thermo Scientific	NA

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact to Herbert B. Schiller (herbert.schiller@helmholtz-munich.de).

Materials availability

Reagents and enzymes were obtained from the commercial sources described in the STAR Methods key resources table.

Data and code availability

- Proteome raw data and MaxQuant processing tables can be downloaded from the PRIDE repository[40] Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Study cohort

Human lung tissues, derived from lung transplantation (donor and ILD recipients), were obtained from the BioArchive of the Comprehensive Pneumology Center Munich (CPC-M). The autoantibody study population consisted of two independent ILD cohorts. Plasma samples of study cohort 1 (Munich cohort) were derived from IPF and CTD-ILD patients from the BioArchive of the CPC-M. Healthy controls were also received from the BioArchive of the CPC-M according to comparable ages as far as possible, although 1:1 matching was not possible the age range was comparable. Controls showed no signs of infection or respiratory symptoms.

Blood samples were either collected in the respective ILD outpatient clinic during routine visits or in the inpatient unit during evaluation for lung transplantation. Diagnosis of IPF was made by high-resolution computed tomography (HRCT), or, if available, histopathological findings from surgical lung biopsies (SLB) according to the official ATS/ERS/JRS/ALAT statement on the diagnosis and management of IPF. Patients with CTD-ILD had a radiographic proven ILD with underlying systemic sclerosis, Sjogren syndrome, rheumatoid arthritis, polyarthritis or undifferentiated CTD. Information about sex, age, and ethnicity as well as some additional clinical data of patients from cohort 1 and cohort 2 is provided in the Tables S3 and S4.

All patients/healthy controls had to be older than 18 years and all gave written informed consent. The study was performed according to the local ethics committee of the Ludwig-Maximilian University Munich (approval numbers 333-10 and 382-10). Patients gave written informed consent to the DZL broad-consent form and the study was approved by the local ethics committee of the Medizinische Hochschule Hannover (2923-2015).

METHOD DETAILS

Patient material processing

For plasma sampling, fresh venous blood was collected in EDTA-coated vacutainer tubes (Sarstedt, Nümbrecht, Germany). After centrifugation, supernatant plasma was separated from blood cells and immediately stored at 80°C degrees in the CPC BioArchive. For cohort 2, Serum samples were obtained from the collaborators in Hannover, and stored at −80°C.

Protein extraction of end-stage lung tissue was performed from the explanted lungs of ILD patients and healthy donors undergoing lung transplantation. To encompass the potentially higher antigenic variation in ILD samples, we prioritized ILD tissue samples for immunoprecipitation (n = 41). In addition, we added non-ILD control samples from healthy tissue (n = 12) to encompass an even broader range of potential antigenic variance. After lung transplantation, the lung tissue is collected in our CPC BioArchive in a standardized way. Lung tissue is cut into pieces of 0.5 cm^3, snap-frozen in liquid nitrogen, and stored at -80°C degrees. For protein extraction, frozen lung tissue pieces were subject to cryo-bead milling in TissueLyser2 (Qiagen). The resulting powder was resuspended in RIPA buffer (50mM Tris HCl pH 7.4, 150 mM NaCl, 1% Triton X100, 0.5% sodium deoxycholate, 1 mM EDTA, 0.1% SDS) supplemented with protease inhibitor and a phosphatase inhibitor cocktail (both Roche; Penzberg, Germany) and assisted by sonication (10 cycles, 30sec power, 30-sec pause) (Bioruptor, Diagenode). Samples were incubated on ice for 30 minutes and undissolved debris was removed by centrifugation for 5 minutes at 18000g. Bicinchoninacid Assay (BCA) (Pierce), was used to analyze protein concentration. For the protein extraction pool, equal amounts of all samples were pooled.

Differential antigen capture assay (DAC)

To identify autoantigen-autoantibody interactions in ILD, we developed the Differential Antigen Capture assay (DAC), based on immunoprecipitation. Protein G agarose-coupled beads (Pierce, Thermo Fisher Scientific, 20397), which can bind up to 11-15mg human IgG per ml of settled resin (50% of slurry), according to manufacturer information, were used to catch antibodies from plasma. Protein extract from ILD and control patient lung tissue was added subsequently to capture antigens, binding to the antibodies on the beads. Immunoprecipitation was performed in triplicates for each patient.

Plasma was thawed on ice and centrifuged at 16.000xg for 5 minutes at 4°C degrees. As described above, antigen protein extraction was derived from end-stage lung tissue from the explanted lungs of 41 ILD patients and 12 healthy donors undergoing lung transplantation, and equal amounts of all samples were pooled. Protein extractions were also thawed on ice.

First, 20µl of Protein G agarose beads were put in 96 well-filter plates (MultiScreenHTS-BV, 1.2 µm, Millipore) and washed with 200µl wash buffer (0.1% IGEPAL, 5% Glycerol, 50 mM Tris pH 7.4, 150 mM NaCl, Roche EDTA free protease inhibitor) by centrifugation at 100g for 1 minute. Beads were then loaded with a mixture of 195µl of wash buffer and 5µl of plasma (Munich cohort)/serum (Hannover cohort) supernatants, and incubated for 1 hour at room temperature, at 900 rpm shaking. Afterwards, the supernatant was removed from the filter plates by centrifugation at 100g for 1 minute and the beads were washed 3 times with 200µl wash buffer.

Antibody-saturated beads were loaded with 150µg of antigen-containing protein extract in 200ul wash buffer and incubated for 1 hour at room temperature shaking at 400 rpm. Filter plates were washed with wash buffer (without protease inhibitor) three times and then three times with PBS to prepare for proteomic digestion.

Mass spectrometry analysis of cohorts 1 and 2

Immunoprecipitation samples were subject to on-bead digest. Beads were incubated in digestion buffer (50 µl 8 M Urea Hepes pH8, 0.5µg LysC (Thermo Scientific, in ABC), 10 mM DTT) for 1 hour at room temperature in a shaker (600 rpm). Afterwards, 0.5µg Trypsin (Thermo Scientific) was added in 200µl 50 mM ABC with 55 mM CAA and beads were again incubated for 1 hour at room temperature gently shaking (600rpm). 96 well filter plates were centrifuged at 100g for 1 minute and digested peptides as flow-through were collected in clean 96 well plates. Filters were washed again with 50µl quenching buffer [2M Urea, 50 mM Thiourea, 2mM Hepes in 50 mM ABC (= 50mM NH4HCO3)] and centrifuged at 100g for 1 minute. Digestion continued overnight at 37°C and 600rpm and was stopped by acidifying the samples to 1% TFA.

Peptides were purified using stage tips containing a poly(styrene-divinylbenzene) copolymer modified with sulfonic acid groups (SDB-RPS).[41,42] Stage tips were first activated with 100µl ACN. For equilibration, we first ran 100µl 30%MeOH, 1% TFA over the stage tips, and then 200µl 0.2% TFA. The samples were loaded in 1 % TFA. Afterward, stage tips were first washed twice with 100µl isopropanol in 1% TFA and then again with 200µl 0.2% TFA. Finally, samples were eluted with 60µl of 5% Ammonia and 80% ACN. Afterwards, samples were evaporated at 30°C degree (Eppendorf Evaporator Plus). The final eluates were dissolved in 6 µl buffer A* (MS-loading buffer) under sonication.

Approximately 1 µg of peptides were separated in one-hour gradients on a 50-cm long (75-µm inner diameter) column packed in-house with ReproSil-Pur C18-AQ 1.9 µm resin (Dr. Maisch GmbH). Reverse-phase chromatography was performed with an EASY-nLC 1000 ultra-high-pressure system (Thermo Fisher Scientific) coupled to a Q-Exactive Mass Spectrometer (Thermo Scientific). Peptides were loaded with buffer A (0.1% (v/v) formic acid) and eluted with a nonlinear 60-min gradient of 5–60% buffer B (0.1% (v/v) formic acid, 80% (v/v) acetonitrile) at a flow rate of 250 nl/min. After each gradient, the column was washed with 95% buffer B and re-equilibrated with buffer A. Column temperature was kept at 50°C by an in-house designed oven with a Peltier element,[43] and operational parameters were monitored in real-time by the SprayQC software (Scheltema & Mann, 2012). MS data were acquired with a shotgun proteomics method, where in each cycle a full scan, providing an overview of the full complement of isotope patterns visible at that particular time point, is followed by up to ten data-dependent MS/MS scans on the most abundant not yet sequenced isotopes (top10 method) (Michalski et al., 2011a). The target value for the full scan MS spectra was 3×10^6 charges in the 300−1,650 *m/z* range with a maximum injection time of 20 ms and a resolution of 70,000 at *m/z* 400. The resulting mass spectra were processed using the MaxQuant software,[44] which enabled label-free protein quantification.[45] Peak lists were searched against the human Uniprot FASTA database (*November 2016*), and a common contaminants database (247 entries) by the Andromeda search engine.[41,46]

Mass spectrometry analysis of the longitudinal cohort

Longitudinal samples were processed using the DAC workflow described for Munich and Hannover cohorts. Digested peptides were measured on a modified trapped ion mobility spectrometry quadrupole time-of-flight mass spectrometer (timsTOF Pro, Bruker Daltonik) with a nano-electrospray ion source (CaptiveSpray, Bruker Daltonik) coupled to EASY-nLC-1200 system (Thermo Fisher Scientific). Peptides were loaded with buffer A (0.1% (v/v) formic acid) at 60°C on a 50 cm column with a diameter of 75 µm packed in-house with ReproSil-Pur C18-AQ 1.9 µm resin (Dr.Maisch GmbH). A linear gradient of 5–30% buffer B (0.1% formic acid and 80% ACN in LC–MS-grade water) was applied to the column for 45 minutes, followed by an increase to 60% for 5 minutes and a 10-minute wash in 95% buffer B. The total length of the gradient reached 65 min. Mass spectrometric analysis was performed in data-dependent (ddaPASEF) mode.[47]

Mass spectrometric raw files acquired on TimsTOF Pro were analyzed with MaxQuant software (version 1.6.7.0). The UniProt FASTA database (2021 release, UP000005640_9606) was used as a reference with a protein-level FDR of 1%. Enzyme specificity was set to trypsin and LysC with a maximum of two allowed missed cleavages. Label-free quantification was performed with the MaxLFQ algorithm, and a minimum ratio count of 1.

QUANTIFICATION AND STATISTICAL ANALYSIS

Single-cell RNA sequencing data analysis

B and plasma cells from publicly available datasets[38,39,48] were analyzed using the Seurat (v4.1.0) pipeline in R (v4.0.5). Cell clusters that expressed high levels of *CD19*, *CD22*, *MS4A1*, *IGHM*, *IGHA1*, *IGHA2*, *IGHG1*, *IGHG2*, *IGHG3*, *JCHAIN*, *CD79A* and *CD79B* were subsetted and the union of variable genes in the 3 datasets (n = 4,619), calculated with *FindVariableFeatures*, was selected for principal component analysis. The datasets were integrated with the Harmony algorithm (v0.1.0) and local neighborhoods were computed on 15 principal components with *FindNeighbors*. The integrated dataset was further clustered with a resolution of 0.3 and *FindClusters*. DE gene analysis for all comparisons was carried out with the *FindAllMarkers* function, the MAST algorithm for *min.pct*=0.20 and *logfcthreshold*=0.25. The top 5 cluster markers were visualized with pheatmap (v1.0.12).

Calculation of cluster proportions

The frequencies of the B cell clusters were calculated per dataset and disease were calculated with the *prop.table()* function in Seurat. The results were visualized with ggplot2 (v3.3.5) in R.

Enrichment of B cell gene signatures

The AUCell package (v1.12.0) was used to test the enrichment of B cell gene signatures in the dataset. Genes with more than 3 counts were filtered and the cell rankings were computed with *AUCell_buildRankings()*. The enrichment was carried out with *AUCell_calcAUC()* for the top 5 % of all genes in the rankings. The gene signature enrichment values or the expression values for genes expressed by at least 10% of cells in the dataset was visualized with the pheatmap package.

Clinical data

Clinical parameters were obtained at the time when the plasma was collected and included demographics (age, gender, smoking status, lung function [forced vital capacity (FVC) (% pred.), FVC (l), and diffusing capacity of the lung for carbon monoxide (DLCO) (SB) (% pred.)]. Transplant-free survival was retrospectively evaluated after plasma sampling.

Statistical analysis

Statistical analysis of clinical data included t-test statistics, ANOVA tests, Fisher's exact test and Kaplan-Meier survival analysis of lung transplantation-free survival using the GraphPad Prism 5 software.

To identify the most commonly and robustly found autoantigens, we used a score based on the sum of the Student's T-test statistics values: ILD patients were tested against control patients and of the significantly enriched proteins (FDR<10%), the score was calculated, per patient per protein (sum score of significantly enriched antigens). Antigens were only included if they were identified in at least three patients in at least one of the four study groups and a cut-off sum score value >4 was used.

All other statistical and bioinformatics operations (such as normalization, data integration, annotation enrichment analysis, hierarchical clustering, and principal component analysis), were run with the Perseus software package (version 1.5.3.0 and 1.6.2.3. and 1.6.10.50).[45]

Proteomics values were filtered, normalized, and imputed using the DEP package (v1.21.0). Visualization and additional statistical analysis of longitudinal samples were completed within the R environment using ggpubr (v0.5.0).