

Mass cytometry identifies the source of cellular and functional changes of gene expression profiles associated with Idiopathic Pulmonary Fibrosis mortality

Qin Li

Yale University

Jiehuan Sun

Yale University

Khadir Raddassi

Yale University

Hongyi Pan

Yale University

Xiting Yan

Yale University

Shelly Ren

Yale University

Hongyu Zhao

Yale University

Argyrios Tzouvelekis

National and Kapodistrian University of Athens

Ruth Montgomery

Yale University

Erica Herzog

Yale University

Naftali Kaminski

Yale University

Jose David Herazo-Maya (✉ josedavidh@hotmail.com)

Research article

Keywords: Idiopathic Pulmonary Fibrosis, Mass Cytometry, Biomarkers, PBMC

Posted Date: August 9th, 2019

DOI: <https://doi.org/10.21203/rs.2.12579/v1>

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

[Read Full License](#)

Abstract

Background We have previously identified a 52-gene signature that distinguishes Idiopathic Pulmonary Fibrosis (IPF) patients with low and high risk genomic profiles with significant differences in mortality. Here, we aimed to apply mass cytometry to characterize phenotypic and functional differences in immune cells between controls and IPF patients with low and high-risk genomic profiles based on the 52-gene signature. **Methods** PBMCs from 20 subjects, including 9 low-risk, 6 high-risk and 5 controls were barcoded with unique palladium-based combinatorial labels and stained as a single multiplexed sample using mass cytometry. 30 antibodies were studied simultaneously. K-means clustering was used to classify cell events into cell subsets. Differences in immune cell subpopulations were examined by generalized linear mixed-effects model. Linear mixed-effects model was used for differential analysis of functional targets in each individual cell subset between groups. **Results** Compared to controls, IPF patients were characterized by significantly increased number of classical monocytes, mDCs, activated Th and Tc cells, and decreased naïve Th cells. We observed a significantly higher frequency of classical monocytes in IPF patients with high versus low risk genomic profiles. Significantly increased levels of CD121b and decreased levels of CD28, LCK, ICOS, CD127, CD96 and ETS1 were observed in monocytes, T-cells, B-cells and granulocytes in high-risk compared to low-risk IPF suggesting these cells as the source of our previously identified gene expression findings. **Conclusions** Our study identified novel immune aberrations in IPF patients and the source of cellular and functional changes of genomic risk profiles associated with IPF mortality.

Background

Idiopathic pulmonary fibrosis (IPF) is a chronic, progressive and irreversible interstitial lung disease of unknown etiology. The median survival of untreated IPF patients is approximately 3.5 years and the clinical course of IPF is unpredictable^[1, 2]. It progresses very rapidly in some patients while others can have a stepwise decline. Another subgroup of patients progress more insidiously with periods of clinical stability^[1-3]. No reliable clinical parameters are available to reflect disease prognosis. Consequently, the identification of candidate biomarkers that can accurately predict disease progression and their cellular source is critical for development of clinical prediction models and for the identification of novel therapeutic targets.

Peripheral blood has been a valuable source for biomarker discovery in IPF. Circulating humoral factors studied in this context include MMP7^[4-6], CXCL-13^[7, 8], IL-8^[9, 10], CCL-18^[11] and anti-HSP70 antibody^[12], levels of which are increased in IPF and highly correlated with disease progression. Changes in distribution patterns of peripheral immune cell subsets, such as CD4⁺CD28⁺T cells^[13, 14], fibrocytes^[15] and Semaphorin7a⁺ regulatory T cell subset^[16], are associated with mortality in IPF. We previously identified a 52-gene expression signature in peripheral blood mononuclear cells (PBMCs) that was associated with transplant-free survival (TFS) in two independent cohorts^[14]. In a follow up study, the 52-gene signature was used to classify IPF into low-risk and high-risk genomic profiles with significant

differences in mortality and TFS [17] findings validated in 6 independent IPF cohorts. In the present study, we hypothesized that genomic risk profiles in PBMC are reflective of changes in the proportion and function of specific immune cell populations in IPF, thus we sought to determine these differences in IPF patients when compared to control subjects and in IPF patients with low versus high-risk gene expression profiles based on selected proteins of the 52-gene signature. To validate our hypothesis, we used mass cytometry by time of flight (CyTOF), a highly parameterized single cell-based platform that can determine functional characteristics of several immune cell subsets in the same subject [18-20]. We found distinct immune cell subpopulation patterns and functional markers on each immune cell subsets in PBMCs that distinguished IPF patients from controls, and high-risk from low-risk IPF patients based on the 52-gene signature. Our study identified novel immune aberrations in IPF patients when compared to controls and the cellular source of gene expression changes associated with high risk of mortality in patients with IPF.

Methods

Human participants

Studies were carried out with explicit informed consent on protocols that were approved by the Human Investigation Committee at Yale University School of Medicine (approval number 1307012431). IPF diagnosis was based on current guidelines of the American Thoracic Society and European Respiratory Society [21]. Patients who were followed at the Yale Interstitial Lung Disease Center of Excellence were eligible to enroll. Healthy, age-matched controls were recruited from Yale's Program on Aging and from the greater New Haven community.

Isolation of peripheral blood mononuclear cells (PBMCs)

PBMCs were isolated from whole blood by using Ficoll paque-based separation (Stem Cell Technologies, Vancouver, BC, Canada) and stored in liquid nitrogen for future studies.

Study strategy and patient population

We designed a CyTOF panel of 30 metal isotope-tagged antibodies to analyze and compare the phenotypic and functional heterogeneity of PBMCs across IPF patients and control (Figure 1 and Table 1). The phenotypic panel contained 15 lineage markers to distinguish the major immune cell populations of PBMCs. The functional panel contained 15 selected protein markers of a 52-gene signature previously associated with increased IPF mortality. These two panels were combined with PBMC samples from control individuals (N=5) and IPF patients (Total IPF N=15; patients with low-risk genomic profile N=9 and patients with high-risk genomic profile N=6). The patient's information is listed in Table 2 and Table 3.

Reagents

Antibodies, manufacturers, and elemental isotope used for labeling cells are listed in Table 1 as described previously [22, 23]. Therein, these antibodies were custom conjugated by Fluidigm, including CD2-155GD,

CD28-166Er, CD96-153Sm, CD121b-163Dy, MCEMP1-143Nd, CXCR6-154Sm, STAT4-161Dy, ETS1-168Er and HAPTOGLOBIN-156Gd. Maxpar® Water (# 201069, Fluidigm, South San Francisco, CA) was used as sample carrier and Maxpar® Cell Staining Buffer (# 201068, Fluidigm, South San Francisco, CA) was used as staining and washing media for PBMCs. Intracellular staining procedure was performed with Foxp3/Transcription Factor Staining Buffer Set (#00-5523, eBioscience, San Diego, CA). The barcoding experiment was carried out using Cell-ID 20-Plex Pd Barcoding Kit (# 201060, Fluidigm, South San Francisco, CA).

Palladium (Pd) barcoding and staining by automated robotic platform

Twenty-samples were processed in batches using a state of the art automated Biomek robotic platform (custom BeckmanCoulter Biomek® robotic machine FXp) for cell barcoding and staining to minimize experimental variations [24, 25]. Samples were washed and counted using automated counter (Guava). 500,000 cells for each sample were transferred to a V bottom 96 well plate and incubated with 50uM of Cisplatin for 1 min to stain dead cells. The plate was centrifuged at 1000x *g* for 5 min, supernatant discarded and washed pellets three-times with PBS. The samples were fixed, washed and suspended with Barcode Perm Buffer. Then, 100 µl of PBS was added to each 20-plex barcoding tube and mixed. 30 µl of the barcoding buffer were transferred to each well followed by mixing well with the sample and incubation at RT for 30 min. The samples were washed three-times before pooling into one eppendorf tube. After washing with Maxpar® Cell Staining Buffer (# 201068, Fluidigm, South San Francisco, CA), the barcoded samples were stained for surface markers and incubated for 30 min at room temperature. PBMCs were then washed, fixed and permeabilized with Foxp3 / Transcription Factor Staining Buffer Set (#00-5523, eBioscience, San Diego, CA), followed by mixing well with intracellular antibodies for 30 min at room temperature. At the end of incubation, the samples were washed and incubated with iridium intercalator (# 201192A, Fluidigm, South San Francisco, CA) diluted into MaxPar® Fix and Perm Buffer (# 201067, Fluidigm, South San Francisco, CA) at 4°C overnight. The next day, cells were washed twice with PBS, and once with Maxpar® Water (# 201069, Fluidigm, South San Francisco, CA) before injection into the CyTOF.

Cytometry by Time-Of-Flight (CyTOF)

Samples were analyzed on a CyTOF 2 instrument (Fluidigm, South San Francisco, CA). Cells were suspended in EQ Four-element calibration beads (# 201078, Fluidigm, South San Francisco, CA) with MaxPar water to a final concentration of 10⁶/mL. Cell suspension was injected at flow rate of 0.045 ml/min after filtration through a 35um nylon mesh (cell strainer cap tubes, BD Biosciences, San Jose, CA). FCS files were generated by CyTOF instrument control soft ware v6.0.622 (Fluidigm, South San Francisco, CA).

Data analysis

Preprocess

The Single Cell Debarcoder was used to debarcode the barcoded FCS file into separate FCS files for each subject [26]. The marker intensities in each sample were log10-transformed. Cell events with low intensities for the two DNA markers (191Ir and 193Ir) and cell events with high intensity for the cell viability marker (195Pt) were gated out using the threshold defined as three median absolute deviations below the median intensity for the pooled cells.

viSNE

viSNE was implemented in MATLAB and Cytobank [27]. 10,000 cells were uniformly subsampled from each group. The phenotypic markers were selected for building up the structure of viSNE map. viSNE maps showed marker expression levels from low to high as color coded. Analysis was performed multiple times with downsampling to optimize graphical output and the results were comparable and stable.

Statistical analysis

To decide whether each cell event belongs to the cell subset of interest, we performed the following analysis: for each phenotypic marker, we classified all cell events into cell events having low, medium, or high expression of this marker using K-means clustering method with the number of clusters fixed at three. Based on the expression levels of all relevant phenotypic markers for a given cell subset of interest, cell events were decided whether they belonged to this cell subset. Because the definitions of cell subsets are not mutually exclusive, a cell event can be classified into multiple cell subsets. For clarity, we excluded ambiguous cell events from our analysis, which were classified into more than one cell subsets. For a given cell subset, we conducted differential analysis in the proportions of this cell subset for both IPF patients versus controls and low-risk versus high-risk patients. The differential analysis was done using a generalized linear mixed-effects model for binomial data, in which the group label was included as a fixed effect and random intercepts were included to account for intra-patient correlations. For functional markers on each of the cell subset, we conducted differential analyses for both IPF patients versus controls and low-risk versus high-risk patients. The differential analysis was done using a linear mixed-effects model, in which the group label (IPF versus controls, IPF low-risk versus IPF high-risk) was included as a fixed effect and random intercepts were included to account for intra-patient correlations. We declared a p-value < 0.05 to be significant. All statistical analyses were conducted using R statistical software (v.3.4.0) and the lme4 package.

Results

Mass cytometry identifies differences in immune cell subpopulations between IPF patients and control subjects.

Our study design is summarized in figure 1. viSNE plots are shown in two dimensions and each dot represents a single cell positioned in the two dimensional space. The intensity of each marker was visualized as color bar indicated (Figure 2). By K-mean clustering, 17 immune cell subpopulations were inferred based on 15 selected phenotypic surface markers. This robust unsupervised clustering algorithm

allowed us to automatically classify the number of cell events into each of the cell subsets studied based on more than ten markers. We compared cell frequencies between IPF patients and control subject on the total of 17 immune cell populations analyzed and identified distinct immune cell composition between these 2 groups (Figure 3A and B). As indicated in Figure 3C, myeloid DCs and classical monocytes were increased in IPF patients versus controls (myeloid DC, $p<0.05$; classical monocytes, $p<0.001$). In addition to monocytes, activated CD4⁺Th cells and activated CD8⁺Tc cells were also significantly increased in IPF patients versus controls (activated Th cells, $p<0.01$; activated Tc cells, $p<0.001$). Naïve CD4⁺Th cells were significantly reduced in IPF patients versus controls ($p<0.01$).

52-gene signature functional markers differentiate IPF patients and controls

To determine cell-subtype specific changes of functional markers of the 52-gene signature, at single-cell resolution, we analyzed the expression of 15 selected targets of this signature in each of the 17 immune cell subsets studied. We found statistically significant differences in several cell subsets between IPF patients and controls. Notably, both innate and adaptive immune cell subsets showed differences between IPF and controls. In non-classical monocytes, IPF patients exhibited lower expression of CD127 (Figure 4A, $p<0.01$) and higher expression of MCEMP1 (Figure 4A, $p<0.05$). Specially, among CD4⁺T cell subsets (helper T cells, Th cells), certain functional markers were found to be significantly different in naïve Th cell, effector Th and memory Th cell subpopulations. Within naïve Th cells, the expression of CD121b, FTL3, MCEMP1, CD96, CXCR6, HAPTOGLOBIN and ICOS, were significantly higher in IPF patients versus controls (Figure 4B, CD121b, $p<0.01$; FLT3, $p<0.01$; MCEMP1, $p<0.05$; CD96, $p<0.05$; CXCR6, $p<0.05$; HAPTOGLOBIN, $p<0.05$; ICOS, $p<0.05$). In effector Th cells, we also observed higher expression of ICOS, FLT3, MCEMP1 and CD121b in IPF patients (Figure 3C, ICOS, $p<0.01$; FLT3, $p<0.01$; MCEMP1, $p<0.01$; CD121b, $p<0.05$). Moreover, STAT4 and ETS1 were increased in IPF patients while CD2 expression was decreased significantly (Figure 4C, STAT4, $p<0.01$; ETS1, $p<0.05$; CD2, $p<0.05$). In memory Th cells, we observed increased expression of CXCR6 and ICOS in IPF patients (Figure 4D, CXCR6, $p<0.05$; ICOS, $p<0.05$). Also, significantly decreased expression of CD47 in IPF patients was found in this subpopulation (Figure 4D, $p<0.05$). Similarly, we also analyzed these 15 functional makers in different CD8⁺T (cytotoxic T cells, Tc cells) cell subpopulations. Compared with controls, STAT4 was significantly increased in naïve Tc cell subpopulation in IPF (Figure 4E, $p<0.05$) while CXCR6 was increased (Figure 4F, $p<0.05$) in activated Tc cells from IPF patients. Memory Tc cell subpopulation in IPF patients showed increased expression of ICOS and HAPTOGLOBIN (Figure 4G, ICOS, $p<0.01$; HAPTOGLOBIN, $p<0.05$). In terms of B cells, we found significant increase of CD127 in naïve B cells (Figure 4H, $p<0.05$) and CD96 in memory B cells from IPF patients (Figure 4I, CD96, $p<0.01$). Differences in innate immune cell subsets were also noted in basophils, granulocytes and mDCs. In basophils, the expression of ICOS was significantly higher in IPF patients versus controls (Figure 4J, $p<0.05$). In granulocytes, the levels of CD235a/b, MCEMP1 and CD96 were higher in IPF patients versus controls (Figure 4K, CD235a/b, $p<0.01$; MCEMP1, $p<0.01$; CD96, $p<0.01$). In mDCs, we observed lower ETS1 (Figure 4L, $p<0.01$) and higher CD96 (Figure 4L, $p<0.05$) in IPF patients versus controls.

Differences in immune cell subpopulations in IPF patients with low-risk and high-risk genomic profiles based on the 52-gene signature.

We studied the immune cell distribution in IPF patients with low-risk and high risk genomic profiles based on a 52-gene signature in order to determine the cellular source of gene expression changes associated with increased IPF mortality. The intensities of individual molecules on each cell plot were present in viSNE plot (Figure 5). The frequencies of cell subsets in IPF patients with low-risk and high-risk genomic profiles based on the 52-gene signature are shown in Figure 6A and B. In general, the subset distribution between PBMC samples from low high-risk patients was similar with the exception of a significant increase in the frequency of classical monocytes in patients with a high-risk genomic profile compared to low risk ($p < 0.05$, Figure 6C). The proportion of naïve Th cells was decreased in high-risk versus low-risk IPF patients (25.7% versus 33.8%), however, this difference was not statistically significant ($p = 0.12$).

52-gene signature functional markers differentiate IPF patients with low-risk versus high-risk genomic profiles and identify the cellular source of gene expression changes

We identified significant differences in functional markers from the 52-gene signature between IPF patients with low-risk versus high-risk genomic profiles. CD121b was significantly higher in classical monocytes in high-risk versus low-risk IPF patients (Figure 7A, $p < 0.05$). In non-classical monocytes, CD127, CD96, and LCK were lower in high-risk versus low-risk IPF patients (Figure 7B, CD96, $p < 0.05$; LCK, $p < 0.05$; CD127, $p < 0.05$). Analysis of the CD4⁺ T helper cell compartments showed that the expression of CD28 was lower in high-risk versus low-risk IPF patients in the memory Th cell subset (Figure 7C, CD28, $p < 0.001$). Lower expression of ETS1 and CD96 was detected in high-risk versus low-risk IPF patients in the effector Tc cell subset (Figure 7D, ETS1, $p < 0.01$; CD96, $p < 0.05$). LCK expression was lower in high-risk versus low risk IPF subjects in the memory Tc cell subset (Figure 7E, LCK, $p < 0.01$). In B cell populations, high expression of ICOS was noted between high-risk versus low-risk IPF patients in naïve B cells (Figure 7F, ICOS, $p < 0.05$). Within IPF patients, significant differences were also noted in granulocytes, where increased levels of MCEMP1 and CD121b were observed in high-risk versus low-risk IPF patients (Figure 7G, MCEMP1, $p < 0.0001$; CD121b, $p < 0.05$).

In summary, we present directional changes of all the functional makers in low-risk versus high-risk IPF patients in Figure S1. Considering that one of the main goals of this study was to identify the cellular source of the gene expression changes of the 52-gene signature in IPF of bulk RNA from PBMC^[14, 17], we further compared cell-type specific changes with our previous gene expression results (Table 4). As indicated, MCEMP1 and CD121b showed increased expression and CD28, CD96, CD127, ETS1, LCK and ETS1 showed decreased expression in both bulk RNA from PBMC in low-risk and high-risk IPF subgroups by gene expression and by mass cytometry analysis independently, validating the cellular source of functional targets of the 52-gene signature.

Discussion

We have previously identified and validated a 52-gene signature in PBMC predictive of mortality and poor disease outcomes in IPF [14, 17]. Here, we performed deep immunophenotyping using mass cytometry in IPF samples and controls, compared the immune composition between these two groups and in IPF patients with high and low genomic profiles based on the 52 gene signature associated with their respective high and low risk of mortality. We also analyzed 15 selected functional markers of the 52-gene signature by mass cytometry to identify the cellular source of gene expression changes and compared the levels of these markers in IPF and controls samples and between IPF patients with different genomic risk profiles.

When we analyzed the immune cell composition of IPF and control samples, we observed that monocytes and myeloid dendritic cells were significantly increased in PBMCs of IPF patients. [28, 29] Our finding demonstrating increased monocytes in IPF patients deserves special attention. Increased monocyte counts appear to have prognostic value in IPF and other fibrotic lung diseases suggesting a role in lung fibrosis progression [29, 30]. In this study, we demonstrated increased proportions of classical monocytes in patients with a high-risk genomic profile based on the 52-gene signature, previously associated with increased IPF mortality [14, 17]. Our findings validate a very large study by Scott et al involving 7000 patients with fibrotic lung diseases from five independent cohorts [30] showing that increased monocyte counts were associated with increased mortality and poor disease outcomes in IPF and other fibrotic lung diseases. While the findings by Scott et al are impressive, they are limited by the fact that authors did not perform detailed immunophenotyping of the monocytes involved in IPF progression. Our study shows that increased proportions of classical monocytes, particularly those expressing CD121b are associated with increased risk of mortality in IPF. CD121b (Interleukin 1 receptor, type), one of the IPF mortality prediction genes when up-regulated interacts with TOLLIP, a gene that encodes a ubiquitin-binding protein involved in interleukin-1 receptor trafficking and in the turnover of IL1R-associated kinase. Interestingly, TOLLIP variants have been associated with IPF survival [31] thus follow up studies are required to establish the relationship between these two proteins in patients with high risk genomic profiles based on the 52-gene signature. It is important to mention that human findings suggesting the role of monocytes in IPF progression are supported by murine studies in bleomycin-induced lung fibrosis. A recent report showed that circulating Ly6C^{hi} monocytes facilitated the progression of pulmonary fibrosis in mice exposed to bleomycin through direction of alternatively activated profibrotic macrophages [32].

Besides monocytes, we also found increased proportions of circulating dendritic cells in IPF patients when compared to control subjects. While previous reports have shown increased immature and mature DCs in the lung tissue [33] and bronchoalveolar lavage [34] from IPF patients, to our knowledge, our report is the first to demonstrate increased levels of DCs in peripheral blood of IPF patients when compared to controls. This finding is relevant since DCs may have a role in lung fibrosis pathogenesis and progression as demonstrated by the large accumulation of mature DC in mice lungs after bleomycin exposure [35] and the fact that antibodies used to block DC maturation or deplete mature DC chemokine receptors attenuate the pathological hallmarks of pulmonary fibrosis in this model [36].

Our study also identified higher frequencies of activated T lymphocytes, mainly activated Th and Tc subsets in PBMCs of IPF patients when compared to controls. Other reports have shown that relative to controls, lung tissue and BAL fluid from patients with IPF are enriched for several population of T lymphocytes^[37, 38]. One of the unexpected findings of our study, which likely resulted from the advantage of deep immunophenotyping using mass cytometry, is that CD4⁺T cell from IPF patients have a skewed composition with an increased proportion of activated Th cells and decreased proportion of naïve Th cells relative to controls. IPF patients in the high risk genomic profile group based on the 52-gene signature also had a trend towards lower proportions of naïve T cells compared to low risk patients suggesting a role of these cells in IPF progression. Studies showing the association between T cells and IPF pathogenesis and poor disease outcomes have shown conflicting results. While others have demonstrated that increased proportions of T cell subsets, including CD25 expressing CD4 cells, and CXCR3 expressing CD4 and CD8 subsets correlated with IPF^[29], we have previously shown that decreased proportion of CD4 T cells in peripheral blood was found to be associated with poor IPF outcomes. Further studies including more samples and additional validation cohorts will be required to finally determine whether lower proportion of naïve T cells and not increased proportion of activated T cells in peripheral blood are associated with IPF progression and mortality.

Our study not only shed light regarding the differences in immune cell composition in IPF patients when compared to controls and between IPF patients with different genomic subphenotypes associated with increased mortality but more importantly, validated our previous findings demonstrating that decreased levels of T-cell costimulatory pathway genes such as CD28 and LCK can be seen in T-cell subsets of IPF patients with a high-risk genomic profile. Mass cytometry also provided some unexpected findings regarding our previously identified survival prediction genes such as decreased levels of LCK in non-classical monocytes and decrease levels of ICOS in B-cells in patients with a high-risk profile. Decreased levels of CD127, CD96 were also seen in non-classical monocytes whereas ETS1 and CD96 were decreased in seen in effector T-cells in the in the IPF high-risk subgroup. Our study also showed other novel findings such as increased expression MCEMP1^[39], mast cell-expressed membrane protein 1, one of the mortality predictive genes in IPF patients when upregulated based on our previous findings, when IPF samples were compared to control samples. Increased MCEMP1 protein levels were seen in IPF in several immune subpopulations such as classical monocytes, naïve Th cell, effector Th cells and granulocytes. These and other findings of our study will require additional validation in larger cohort of patients.

Despite our impressive results and the validation of our previous gene expression findings, our study obviously has limitations, the two most important being our small sample size and the lack of a validation cohort. While we were able to perform genomic profiling of N=425 individual IPF patients in our 52-gene signature validation study^[17], only a limited number of samples had viable PBMC collected simultaneously for mass cytometry. This limitation may have impaired our ability to detect more granular differences in the proportions of immune cell subpopulations studied across subgroups. It may have also impaired our ability to detect more impressive changes of functional markers in patients with low- and

high-risk genomic profiles. Additional studies with larger number of samples and a validation cohort with simultaneous gene expression and mass cytometry measurements are required to further validate our findings. Future studies should also include the collection of time course samples to determine whether cellular and functional changes are static or dynamic, particularly in response to anti-fibrotic therapies. Another limitation of our study is the purity of PBMC samples, while Ficoll-Paque isolation provides a highly cell purified pellet containing more than 95% mononuclear cells, they may include up to $3\pm 2\%$ granulocytes. In our data, we were able to detect a strong signal from granulocytes between studied groups and although they accounted for only a minor percentage of cells this may be off relevance since a recent study showed that accumulation of aged neutrophils in the lung may cause fibrotic interstitial lung disease^[40] thus, additional studies looking a gene expression and functional differences in neutrophils in IPF is warranted, however, this is beyond the scope of our study.

Conclusions

Our study concludes than IPF patients have increased proportion of classical monocytes, myeloid dendritic cells, activated Th cells and activated Tc cells and decrease proportion of naïve Th cells when compared to healthy controls. Dysregulation of functional protein markers of the 52-gene signature were found in IPF patients compared to controls in several cell subpopulations, the majority of which have not been previously characterized in IPF. Increased proportion of classical monocytes and a trend towards lower naïve Th cells were identified in high-risk versus low-risk IPF patients suggesting that an imbalance between monocytes and naïve T cells may be associated with IPF progression. Finally, we identified the cellular source of changes of protein markers of the 52-gene signature which predominantly happened in classical and non-classical monocytes, memory Th cells, memory Tc cells, effector Tc cells, B cells and granulocytes validating our previous gene expression findings.

Abbreviations

IPF: Idiopathic Pulmonary Fibrosis

CyTOF: Cytometry by time of flight

PBMC: Peripheral blood mononuclear cells

Th: Helper T cells

Tc: Cytotoxic T cells

Declarations

Ethics approval and consent to participate

The study was approved by the Yale University Human Investigation Committee (approval number 1307012431), and written informed consent was obtained from all the participating patients.

Consent for publication

Not applicable.

Availability of data and material

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

Competing interests

Dr. Kaminski reports personal fees from Biogen Idec, Boehringer Ingelheim, Third Rock, Miragen, Pliant, personal fees from Samumed, NuMedii, Indaloo, Theravance, LifeMax, Three Lake Partners, outside the submitted work; In addition, Dr. Kaminski has a patents on New Therapies in Pulmonary Fibrosis Peripheral Blood Gene Expression in IPF. Other authors declare that they have no competing interests.

Funding

JHD is funded by the Harold Amos Medical Faculty Development Program of the Robert Wood Johnson Foundation and the Pulmonary Fibrosis Foundation. NK is funded by NIH grants R01HL141852, R01HL127349, UH3HL123886, U01HL122626, UG3TR002445 and Department of Defence grant PR151124.

Authors' contributions

QL performed experiments, analyzed data and generated figures, wrote the manuscript; JHS performed analyzed data, generated figures and revised manuscript; KR performed experiments; HYP recruited subjects and collected PBMCs; XTY supervised statistical analysis; SR performed CyTOF machine and collected data; HYZ supervised and critiqued analytical approaches; RRM supervised CyTOF sample procurement, data collection; AT , participated in data collection, analysis and revision of the manuscript, ELH characterized patients, supervised PBMC sample procurement, data collection and analysis; NK conceived and designed the project, supervised data collection and analysis, revised the manuscript; JHD conceived and designed the project, supervised data collection, participated in analysis and figure generation and revised the manuscript. All authors approved the final version.

Acknowledgements

Not applicable.

References

1. Martinez FJ, Chisholm A, Collard HR, Flaherty KR, Myers J, Raghu G, Walsh SL, White ES, Richeldi L: **The diagnosis of idiopathic pulmonary fibrosis: current and future approaches.** *Lancet Respir Med* 2017, **5**(1):61-71.
2. Martinez FJ, Collard HR, Pardo A, Raghu G, Richeldi L, Selman M, Swigris JJ, Taniguchi H, Wells AU: **Idiopathic pulmonary fibrosis.** *Nat Rev Dis Primers* 2017, **3**:17074.
3. King CS, Nathan SD: **Idiopathic pulmonary fibrosis: effects and optimal management of comorbidities.** *Lancet Respir Med* 2017, **5**(1):72-84.
4. Rosas IO, Richards TJ, Konishi K, Zhang Y, Gibson K, Lokshin AE, Lindell KO, Cisneros J, Macdonald SD, Pardo A *et al*: **MMP1 and MMP7 as potential peripheral blood biomarkers in idiopathic pulmonary fibrosis.** *PLoS Med* 2008, **5**(4):e93.
5. Richards TJ, Kaminski N, Baribaud F, Flavin S, Brodmerkel C, Horowitz D, Li K, Choi J, Vuga LJ, Lindell KO *et al*: **Peripheral blood proteins predict mortality in idiopathic pulmonary fibrosis.** *Am J Respir Crit Care Med* 2012, **185**(1):67-76.
6. Tzouveleakis A, Herazo-Maya JD, Slade M, Chu JH, Deiuliis G, Ryu C, Li Q, Sakamoto K, Ibarra G, Pan H *et al*: **Validation of the prognostic value of MMP-7 in idiopathic pulmonary fibrosis.** *Respirology* 2017, **22**(3):486-493.
7. Vuga LJ, Tedrow JR, Pandit KV, Tan J, Kass DJ, Xue J, Chandra D, Leader JK, Gibson KF, Kaminski N *et al*: **C-X-C motif chemokine 13 (CXCL13) is a prognostic biomarker of idiopathic pulmonary fibrosis.** *Am J Respir Crit Care Med* 2014, **189**(8):966-974.
8. DePianto DJ, Chandriani S, Abbas AR, Jia G, N'Diaye EN, Caplazi P, Kauder SE, Biswas S, Karnik SK, Ha C *et al*: **Heterogeneous gene expression signatures correspond to distinct lung pathologies and biomarkers of disease severity in idiopathic pulmonary fibrosis.** *Thorax* 2015, **70**(1):48-56.
9. Ziora D, Jastrzebski D, Adamek M, Czuba Z, Kozielski JJ, Grzanka A, Kasperska-Zajac A: **Circulating concentration of markers of angiogenic activity in patients with sarcoidosis and idiopathic pulmonary fibrosis.** *BMC Pulm Med* 2015, **15**:113.
10. Papis SA, Tomos IP, Karakatsani A, Spathis A, Korbila I, Analitis A, Kolilekas L, Kagouridis K, Loukides S, Karakitsos P *et al*: **High levels of IL-6 and IL-8 characterize early-on idiopathic pulmonary fibrosis acute exacerbations.** *Cytokine* 2018, **102**:168-172.
11. Prasse A, Probst C, Bargagli E, Zissel G, Toews GB, Flaherty KR, Olschewski M, Rottoli P, Muller-Quernheim J: **Serum CC-chemokine ligand 18 concentration predicts outcome in idiopathic pulmonary fibrosis.** *Am J Respir Crit Care Med* 2009, **179**(8):717-723.
12. Kahloon RA, Xue J, Bhargava A, Csizmadia E, Otterbein L, Kass DJ, Bon J, Soejima M, Levesque MC, Lindell KO *et al*: **Patients with idiopathic pulmonary fibrosis with antibodies to heat shock protein 70 have poor prognoses.** *Am J Respir Crit Care Med* 2013, **187**(7):768-775.
13. Gilani SR, Vuga LJ, Lindell KO, Gibson KF, Xue J, Kaminski N, Valentine VG, Lindsay EK, George MP, Steele C *et al*: **CD28 down-regulation on circulating CD4 T-cells is associated with poor prognoses of patients with idiopathic pulmonary fibrosis.** *PLoS One* 2010, **5**(1):e8959.

14. Herazo-Maya JD, Noth I, Duncan SR, Kim S, Ma SF, Tseng GC, Feingold E, Juan-Guardela BM, Richards TJ, Lussier Y *et al*: **Peripheral blood mononuclear cell gene expression profiles predict poor outcome in idiopathic pulmonary fibrosis.** *Sci Transl Med* 2013, **5**(205):205ra136.
15. Moeller A, Gilpin SE, Ask K, Cox G, Cook D, Gauldie J, Margetts PJ, Farkas L, Dobranowski J, Boylan C *et al*: **Circulating fibrocytes are an indicator of poor prognosis in idiopathic pulmonary fibrosis.** *Am J Respir Crit Care Med* 2009, **179**(7):588-594.
16. Reilkoff RA, Peng H, Murray LA, Peng X, Russell T, Montgomery R, Feghali-Bostwick C, Shaw A, Homer RJ, Gulati M *et al*: **Semaphorin 7a+ regulatory T cells are associated with progressive idiopathic pulmonary fibrosis and are implicated in transforming growth factor-beta1-induced pulmonary fibrosis.** *Am J Respir Crit Care Med* 2013, **187**(2):180-188.
17. Herazo-Maya JD, Sun J, Molyneaux PL, Li Q, Villalba JA, Tzouveleakis A, Lynn H, Juan-Guardela BM, Risquez C, Osorio JC *et al*: **Validation of a 52-gene risk profile for outcome prediction in patients with idiopathic pulmonary fibrosis: an international, multicentre, cohort study.** *Lancet Respir Med* 2017, **5**(11):857-868.
18. Spitzer MH, Nolan GP: **Mass Cytometry: Single Cells, Many Features.** *Cell* 2016, **165**(4):780-791.
19. Simoni Y, Chng MHY, Li S, Fehlings M, Newell EW: **Mass cytometry: a powerful tool for dissecting the immune landscape.** *Curr Opin Immunol* 2018, **51**:187-196.
20. Brodie TM, Tosevski V: **High-Dimensional Single-Cell Analysis with Mass Cytometry.** *Curr Protoc Immunol* 2017, **118**:5 11 11-15 11 25.
21. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, Behr J, Cottin V, Danoff SK, Morell F *et al*: **Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline.** *Am J Respir Crit Care Med* 2018, **198**(5):e44-e68.
22. Yao Y, Liu R, Shin MS, Trentalange M, Allore H, Nassar A, Kang I, Pober JS, Montgomery RR: **CyTOF supports efficient detection of immune cell subsets from small samples.** *J Immunol Methods* 2014, **415**:1-5.
23. Yao Y, Welp T, Liu Q, Niu N, Wang X, Britto CJ, Krishnaswamy S, Chupp GL, Montgomery RR: **Multiparameter Single Cell Profiling of Airway Inflammatory Cells.** *Cytometry B Clin Cytom* 2017, **92**(1):12-20.
24. Nassar AF, Wisnewski AV, Raddassi K: **Automation of sample preparation for mass cytometry barcoding in support of clinical research: protocol optimization.** *Anal Bioanal Chem* 2017, **409**(9):2363-2372.
25. Nassar AF, Wisnewski AV, Raddassi K: **Progress in automation of mass cytometry barcoding for drug development.** *Bioanalysis* 2016, **8**(14):1429-1435.
26. Zunder ER, Finck R, Behbehani GK, Amir el AD, Krishnaswamy S, Gonzalez VD, Lorang CG, Bjornson Z, Spitzer MH, Bodenmiller B *et al*: **Palladium-based mass tag cell barcoding with a doublet-filtering scheme and single-cell deconvolution algorithm.** *Nat Protoc* 2015, **10**(2):316-333.
27. Amir el AD, Davis KL, Tadmor MD, Simonds EF, Levine JH, Bendall SC, Shenfeld DK, Krishnaswamy S, Nolan GP, Pe'er D: **viSNE enables visualization of high dimensional single-cell data and reveals**

- phenotypic heterogeneity of leukemia.** *Nat Biotechnol* 2013, **31**(6):545-552.
28. Feghali-Bostwick CA, Tsai CG, Valentine VG, Kantrow S, Stoner MW, Pilewski JM, Gadgil A, George MP, Gibson KF, Choi AM *et al*: **Cellular and humoral autoreactivity in idiopathic pulmonary fibrosis.** *J Immunol* 2007, **179**(4):2592-2599.
29. Moore BB, Fry C, Zhou Y, Murray S, Han MK, Martinez FJ, Flaherty KR, The CI: **Inflammatory leukocyte phenotypes correlate with disease progression in idiopathic pulmonary fibrosis.** *Front Med* 2014, **1**(56).
30. Scott MKD, Quinn K, Li Q, Carroll R, Warsinske H, Vallania F, Chen S, Carns MA, Aren K, Sun J *et al*: **Increased monocyte count as a cellular biomarker for poor outcomes in fibrotic diseases: a retrospective, multicentre cohort study.** *Lancet Respir Med* 2019, **7**(6):497-508.
31. Noth I, Zhang Y, Ma SF, Flores C, Barber M, Huang Y, Broderick SM, Wade MS, Hysi P, Scurba J *et al*: **Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study.** *Lancet Respir Med* 2013, **1**(4):309-317.
32. Gibbons MA, MacKinnon AC, Ramachandran P, Dhaliwal K, Duffin R, Phythian-Adams AT, van Rooijen N, Haslett C, Howie SE, Simpson AJ *et al*: **Ly6Chi monocytes direct alternatively activated profibrotic macrophage regulation of lung fibrosis.** *Am J Respir Crit Care Med* 2011, **184**(5):569-581.
33. Marchal-Somme J, Uzunhan Y, Marchand-Adam S, Valeyre D, Soumelis V, Crestani B, Soler P: **Cutting edge: nonproliferating mature immune cells form a novel type of organized lymphoid structure in idiopathic pulmonary fibrosis.** *J Immunol* 2006, **176**(10):5735-5739.
34. Tsoumakidou M, Karagiannis KP, Bouloukaki I, Zakyntinos S, Tzanakis N, Siafakas NM: **Increased bronchoalveolar lavage fluid CD1c expressing dendritic cells in idiopathic pulmonary fibrosis.** *Respiration* 2009, **78**(4):446-452.
35. Bantsimba-Malanda C, Marchal-Somme J, Goven D, Freynet O, Michel L, Crestani B, Soler P: **A role for dendritic cells in bleomycin-induced pulmonary fibrosis in mice?** *Am J Respir Crit Care Med* 2010, **182**(3):385-395.
36. Trujillo G, Hartigan AJ, Hogaboam CM: **T regulatory cells and attenuated bleomycin-induced fibrosis in lungs of CCR7^{-/-} mice.** *Fibrogenesis Tissue Repair* 2010, **3**:18.
37. Nuovo GJ, Hagood JS, Magro CM, Chin N, Kapil R, Davis L, Marsh CB, Folcik VA: **The distribution of immunomodulatory cells in the lungs of patients with idiopathic pulmonary fibrosis.** *Mod Pathol* 2012, **25**(3):416-433.
38. Fireman E, Vardinon N, Burke M, Spizer S, Levin S, Endler A, Stav D, Topilsky M, Mann A, Schwarz Y *et al*: **Predictive value of response to treatment of T-lymphocyte subpopulations in idiopathic pulmonary fibrosis.** *Eur Respir J* 1998, **11**(3):706-711.
39. Li K, Wang SW, Li Y, Martin RE, Li L, Lu M, Lamhamedi-Cherradi SE, Hu G, Demissie-Sanders S, Zheng J *et al*: **Identification and expression of a new type II transmembrane protein in human mast cells.** *Genomics* 2005, **86**(1):68-75.
40. Kim JH, Podstawka J, Lou Y, Li L, Lee EKS, Divangahi M, Petri B, Jirik FR, Kelly MM, Yipp BG: **Aged polymorphonuclear leukocytes cause fibrotic interstitial lung disease in the absence of regulation by**

Tables

Table 1. CyTOF antibody panel

Antigen Target	Clone Number	Supplier	Elemental Isotope
CD3	SP34-2	Fluidigm	170Er
CD4	RPA-T4	Fluidigm	145Nd
CD8	RPA-T8	Fluidigm	146Nd
CD14	RMO52	Fluidigm	160Gd
CD19	HIB19	Fluidigm	142Nd
CD20	2H7	Fluidigm	147Sm
CD11c	Bu15	Fluidigm	159Tb
CD16	3G8	Fluidigm	165Ho
CD27	O323	Fluidigm	167Er
CD38	HIT2	Fluidigm	172Yb
CD66a	CD66a-B1.1	Fluidigm	149Sm
CD123	6H6	Fluidigm	151Eu
CD45	HI30	Fluidigm	89Y
CD45RA	HI100	Fluidigm	169Tm
HLA-DR	L243	Fluidigm	174Yb
CD2	TS1/8	Fluidigm	155Gd
CD28	CD28.2	Fluidigm	166Er
CD47	CC2C6	Fluidigm	209Bi
CD96	NK92.39	Fluidigm	152Sm
CD121b	MNC2	Fluidigm	163Dy
CD235a/b	HIR2	Fluidigm	141Pr
MCEMP1	Polyclonal sheep IgG	Fluidigm	143Nd
FLT-3	BV10A4H2	Fluidigm	158Gd
ICOS	C398.4A	Fluidigm	148Nd
CXCR6	K041E5	Fluidigm	154Sm
CD127	C398.4A	Fluidigm	171Yb
LCK	LCK-01	Fluidigm	153Eu

STAT4	2H9L5	Fluidigm	161Dy
ETS1	Polyclonal rabbit IgG	Fluidigm	168Er
Haptoglobin	RB21463	Fluidigm	156Gd

Table 2. Demographic characteristics of IPF patients and control

Characteristic	IPF patients (N=15)	Control subjects (N=5)
Age at enrollment, Mean \pm SD	69.4 \pm 4.4	73.8 \pm 5.5
Gender, n		
Males	13	3
Females	2	2
Race, n		
Caucasian	15	5
Black	0	0
Hispanic	0	0
Other	0	0
Smoking status, n		
Ever smoker	12	0
Never smoker	3	5
Spirometry, mean \pm SD		
FVC (%)	74.3 \pm 15.6	-
DLCO (%)	40.2 \pm 12.4	-
FEV1 (%)	78.5 \pm 21.7	-
GAP Index, Mean \pm SD	4.2 \pm 1.6	-
Diagnosis, n		
HRCT+ UIP biopsy	1	-
HRCT	14	-

Table 3. Demographic characteristics of the IPF patients in the two risk groups

Characteristic	Low-risk (N=9)	High-risk (N=6)
Age at enrollment, Mean \pm SD	69.0 \pm 4.3	70 \pm 4.9
Gender, n		
Males	7	6
Females	2	0
Race, n		
Caucasian	9	6
Black	0	0
Hispanic	0	0
Other	0	0
Smoking status, n		
Ever smoker	7	5
Never smoker	2	1
Spirometry, mean \pm SD		
FVC (%)	78.6 \pm 12.2	68.0 \pm 19.0
DLCO (%)	43.2 \pm 13.3	35.7 \pm 10.5
FEV1 (%)	79.1 \pm 22.2	77.7 \pm 23.0
GAP Index, Mean \pm SD	3.9 \pm 1.7	4.8 \pm 1.5
Diagnosis, n		
HRCT+ UIP biopsy	1	0
HRCT	8	6

Table 4. Functional markers with consistent changes from both genetic and protein level

Functional markers	Genetic change (microarray)	Protein change (CyTOF)
	'low-risk' vs. 'high-risk'	'low-risk' vs. 'high-risk'
MCEMP1	↑ in bulk RNA of PBMCs	↑ in granulocytes
CD121b (IL1R2)	↑ in bulk RNA of PBMCs	↑ in granulocytes
		↑ in classical monocytes
CD28	↓ in bulk RNA of PBMCs	↓ in memory Th cells
CD96	↓ in bulk RNA of PBMCs	↓ in effector Tc cells
		↓ in non-classical monocytes
CD127 (IL7R2)	↓ in bulk RNA of PBMCs	↓ in non-classical monocytes
ETS1	↓ in bulk RNA of PBMCs	↓ in effector Tc cells
LCK	↓ in bulk RNA of PBMCs	↓ in memory Tc cells
		↓ in non-classical monocytes

Figures

Figure 1

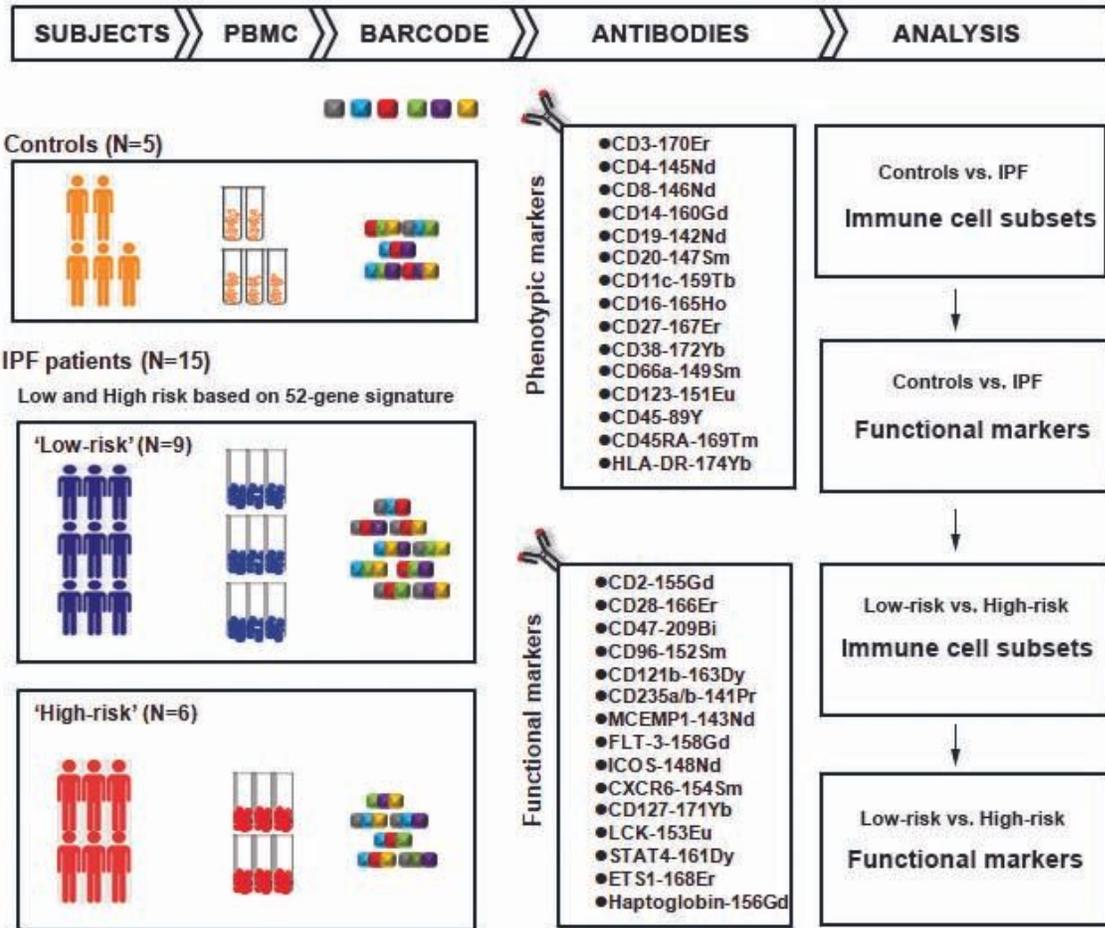


Figure 1

Experiment workflow. The outline summarizes the subjects, antibodies, the experiments performed in each group, and the analysis used.

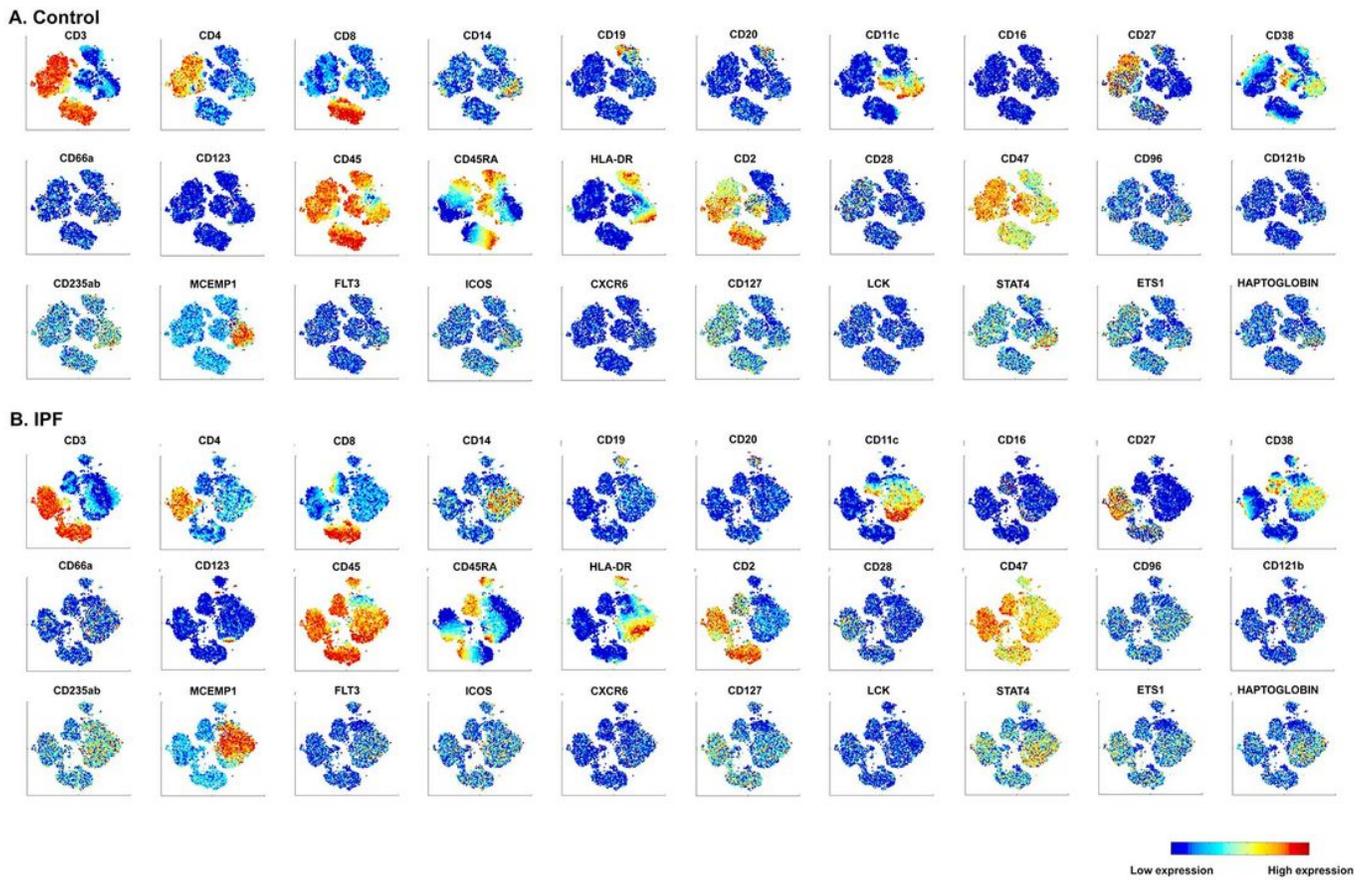


Figure 2

viSNE map of PBMCs from control and IPF. Ten thousand cells from control (A) and IPF (B) were shown in each map. Each point in the viSNE map represents an individual cell, colored by the expression of levels of the maker in the panel's title.

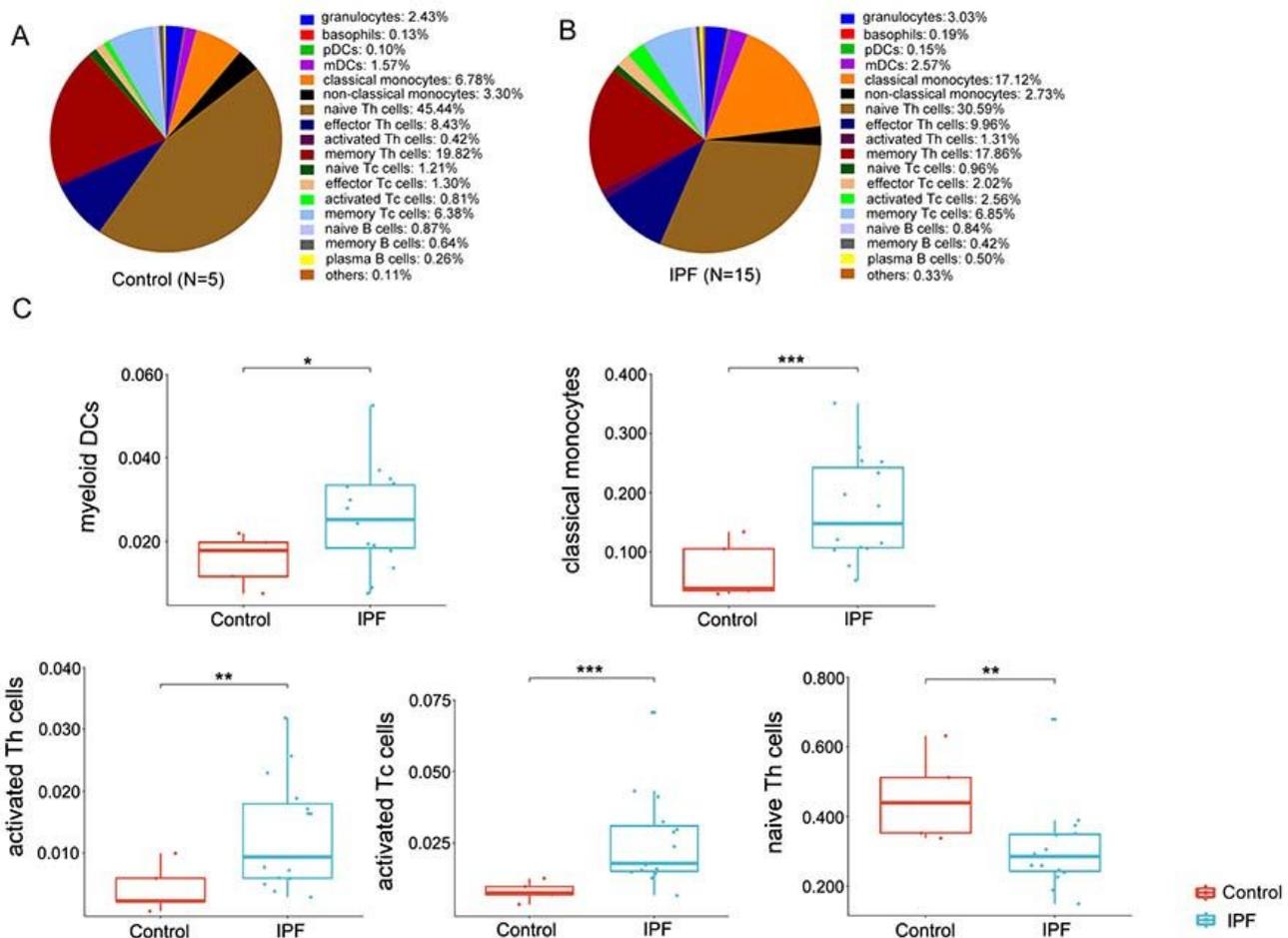


Figure 3

Differences in immune cell proportions between IPF and controls. (A) Pie charts displayed the frequencies of granulocytes, basophils, plasmacytoid dendritic cells (pDCs), myeloid dendritic cells (mDCs), classical monocytes, non-classical monocytes, naïve CD4+T cells (naïve Th cells), effector CD4+T cells (effector Th cells), activated CD4+T cells (activated Th cells), memory CD4+T cells (memory Th cells), naïve CD8+T cells (naïve Tc cells), effector CD8+T cells (effector Tc cells), activated CD8+T cells (activated Tc cells), memory CD8+T cells (memory Tc cells), naïve B cells, memory B cells, and plasma B cells in peripheral blood from 5 normal control subjects (B) and 15 IPF patients (C). (D) Five of 17 subpopulations (mDCs, classical monocytes, activated Th cells, activated Tc cells and naïve Th cells) in IPF patients showed significant changes compared to normal subjects. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$

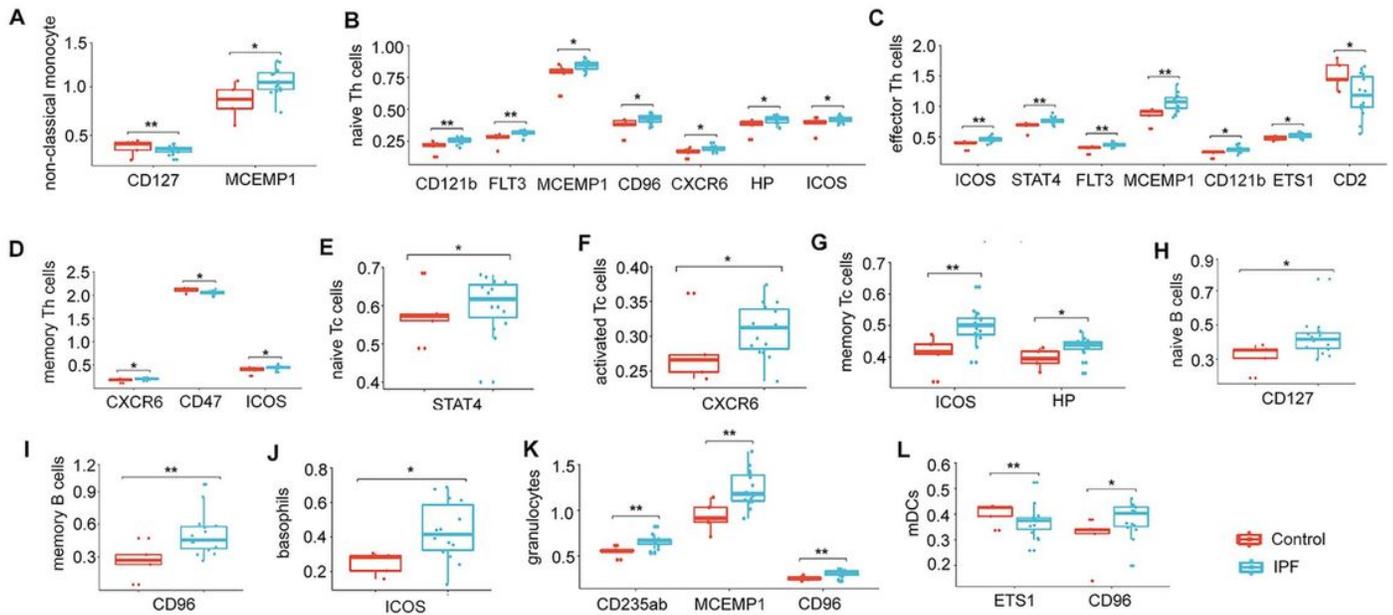


Figure 4

The expression of cell type-specific functional molecules differentiates IPF patients from control subjects. Functional molecules displayed significantly differential changes between IPF patients and normal subjects in non-classical monocytes (A), naïve Th cells (B), effector Th cells (C), memory Th cells (D), naïve Tc cells (E), activated Tc cells (F), memory Tc cells (G), naïve B cells (H) and memory B cells (I), basophils (J), granulocytes (K), mDC (L). . * represents $p < 0.05$, ** represents $p < 0.01$

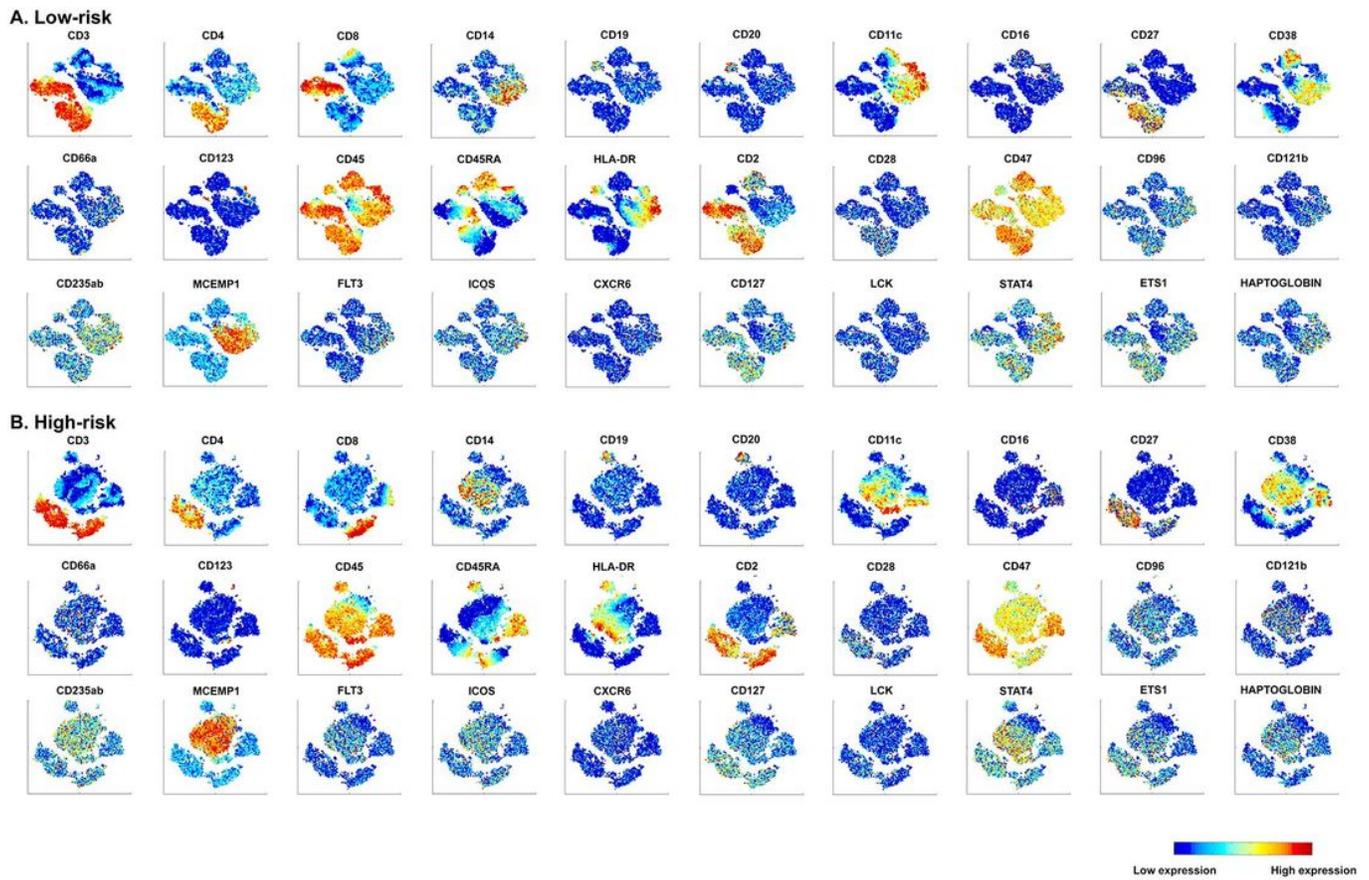


Figure 5

viSNE map of PBMCs from ‘low-risk’ and ‘high-risk’ IPF patients. Ten thousand cells from low-risk (A) and high-risk IPF patients (B) were shown in each map. Each point in the viSNE map represents an individual cell, colored by the expression of levels of the marker in the panel’s title.

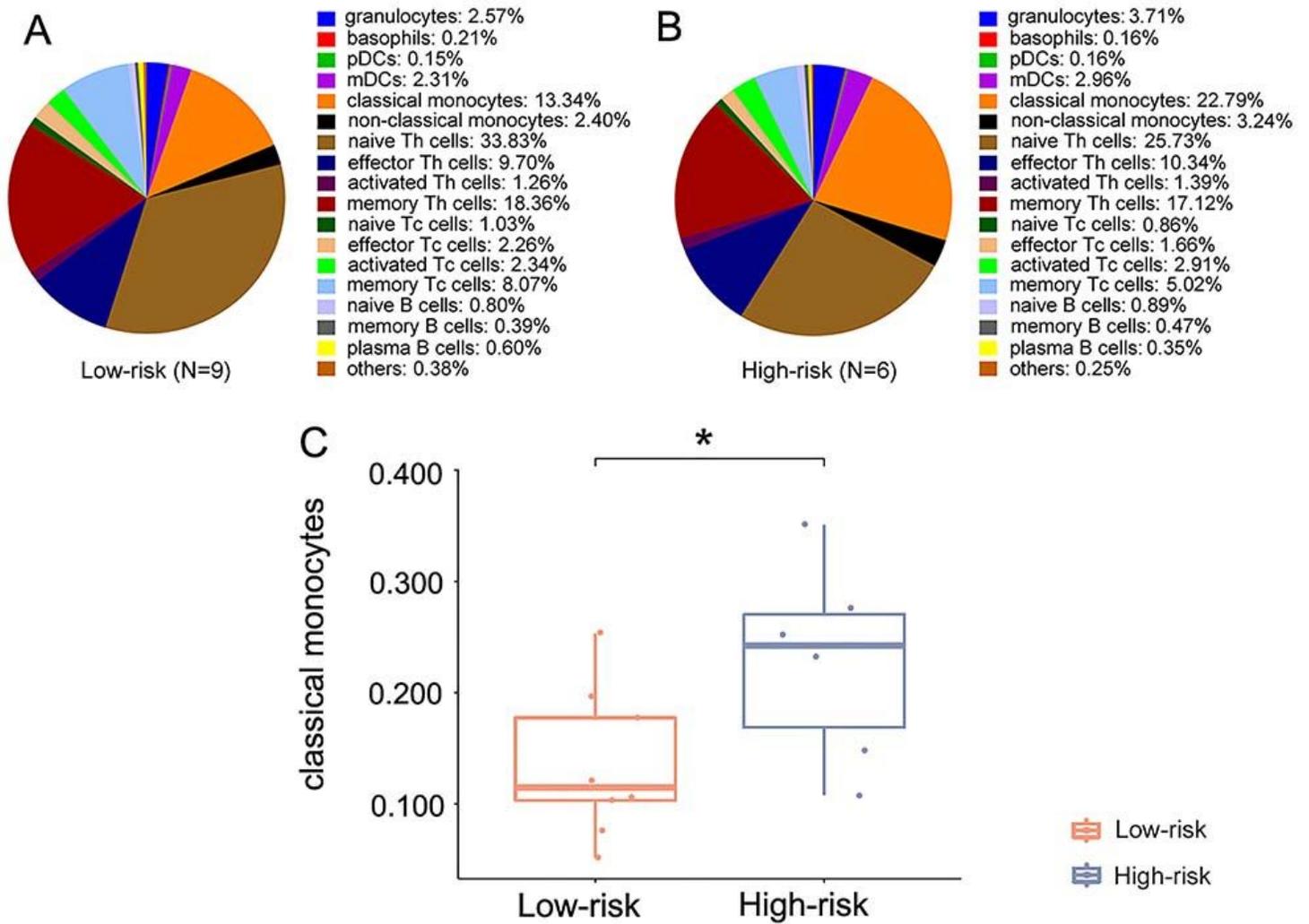


Figure 6

Differences in immune cell proportions between high-risk and low-risk IPF patients. (A) Pie charts displayed the frequencies of granulocytes, basophils, plasmacytoid dendritic cells (pDCs), myeloid dendritic cells (mDCs), classical monocytes, non-classical monocytes, naïve CD4+T cells (naïve Th cells), effector CD4+T cells (effector Th cells), activated CD4+T cells (activated Th cells), memory CD4+T cells (memory Th cells), naïve CD8+T cells (naive Tc cells), effector CD8+T cells (effector Tc cells), activated CD8+T cells (activated Tc cells), memory CD8+T cells (memory Tc cells), naïve B cells, memory B cells, and plasma B cells in peripheral blood from 6 'low-risk' patients (B) and 9 'high-risk' patients (C). (D). The frequency of classical monocytes showed significant difference between 'low-risk' and 'high-risk' patients. * represents $p < 0.05$

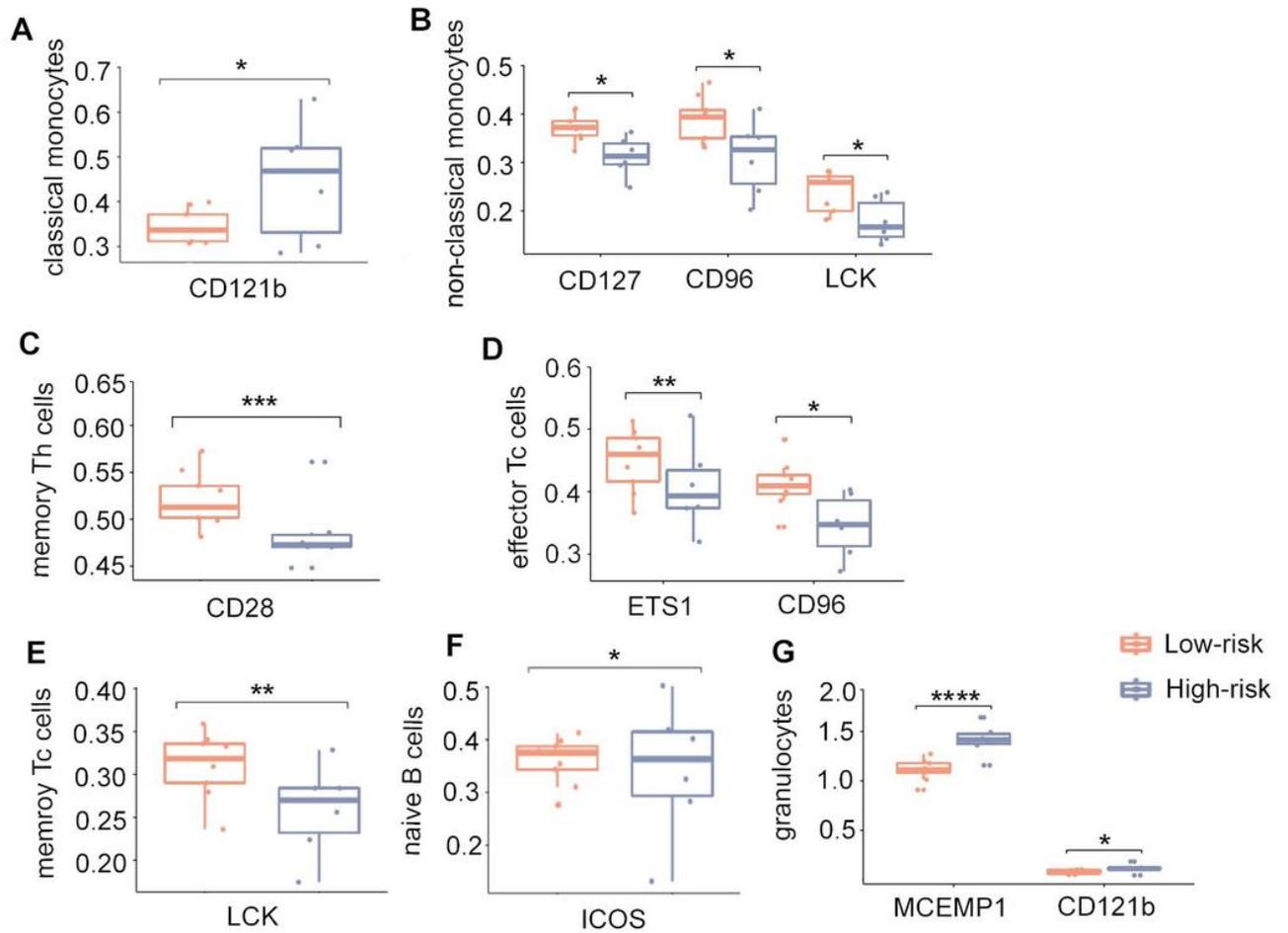


Figure 7

The expression of cell type-specific functional molecules differentiates high-risk from low-risk IPF patients. Functional molecules displayed significantly differential changes between high-risk and low-risk IPF patients classical monocytes (A), non-classical monocytes (B), memory Th cells (C), effector Tc cells (D), memory Tc cells (E), naïve B cell (F) and granulocytes (G). * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$, **** represents $p < 0.0001$

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFileFigureS1.tif](#)