A validated hypoxia-inducible factor (HIF) signature across tissue 1

compartments predicts clinical outcome in human lung fibrosis 2

3

4 Yilu Zhou^{1,2}, Rob M. Ewing^{1,2}, Donna E. Davies^{2,3,4}, Yihua Wang^{1,2,4,5,†}, Mark G. Jones^{2,3,4,5,†}

5

6 ¹Biological Sciences, Faculty of Environmental and Life Sciences, University of Southampton 7 SO17 1BJ, UK. ²Institute for Life Sciences, University of Southampton, Southampton SO17 1BJ, UK. ³Clinical and Experimental Sciences, Faculty of Medicine, University of 8 9 Southampton SO16 6YD, UK. ⁴NIHR Respiratory Biomedical Research Centre, University 10 Hospital Southampton SO16 6YD, UK. ⁵Correspondence should be addressed to M.G.J. (e-11 mail: mark.jones@soton.ac.uk) or Y.W. (e-mail: yihua.wang@soton.ac.uk). [†]Senior authors 12 contributed equally to this work

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

1 Abstract

2 We previously reported that oxidative stress drives pseudohypoxic hypoxia-inducible factor 3 (HIF) pathway activation to promote pathogenetic collagen structure-function in human lung 4 fibrosis (Brereton et al., 2022). Here, through bioinformatic studies we investigate HIF 5 pathway activation status in patients with idiopathic pulmonary fibrosis (IPF) and whether this 6 has prognostic significance. Applying a well-established HIF gene expression signature, we 7 classified publicly available datasets into HIF score-high and score-low groups across multiple 8 tissue compartments. The HIF scores in lung tissue, bronchoalveolar lavage (BAL) and 9 peripheral blood mononuclear cells (PBMC) were increased in IPF patients and significantly 10 correlated with an oxidative stress signature consistent with pseudohypoxic HIF pathway 11 activation. A high HIF score in BAL and in PBMC was a strong independent predictor of 12 mortality in multivariate analysis. Thus, a validated HIF gene signature predicts survival across 13 tissue compartments in IPF and merits prospective study as a non-invasive biomarker of lung 14 fibrosis progression.

1 Introduction

2 We previously reported that altered collagen nano-architecture is a defining feature of 3 idiopathic pulmonary fibrosis (IPF) that dysregulates extracellular matrix (ECM) structure-4 function to promote progressive lung fibrosis (1). We have recently extended these 5 observations, identifying that hypoxia-inducible factor (HIF) pathway activation is required to 6 promote pathologic pyridinlone collagen crosslinking and tissue stiffness by disproportionate 7 induction of collagen-modifying enzymes relative to TGFβ-induced collagen fibril synthesis 8 (2). Furthermore, we identified that this may occur via oxygen-independent (pseudohypoxic) 9 mechanisms, including a decrease in factor inhibiting HIF (FIH) activity due to oxidative stress 10 (2). Oxidative stress was increased in subpopulations of IPF fibroblasts whilst FIH activity 11 was significantly reduced in fibroblasts from patients with lung fibrosis resulting in HIF 12 activation under normoxic conditions. To further assess for HIF activity we applied a validated 13 HIF/hypoxia metagene signature, identifying that HIF activity was increased at sites of active 14 fibrogenesis in IPF tissue as well as within IPF mesenchymal cell populations. Intriguingly, in 15 a small lung mesenchymal stromal RNAseq data set increased HIF pathway activation assessed 16 via gene set variation analysis (GSVA) of this HIF signature was associated with disease 17 progression. Here we sought to further investigate the potential utility of this validated HIF 18 signature as a predictor of disease progression.

Results 1

2 A validated HIF score is increased across tissue compartments in patients with IPF.

3 We applied a 15-gene expression signature (ACOT7, ADM, ALDOA, CDKN3, ENO1, LDHA, MIF, MRPS17, NDRG1, P4HA1, PGAM1, SLC2A1, TP11, TUBB6 and VEGFA), which has 4 5 been validated across multiple populations (3-6) and which enables classification of HIF 6 activity by calculating a HIF score for each sample using GSVA, to publicly available datasets 7 (7). We firstly confirmed that the HIF score was significantly increased in the transcriptome of 8 bulk lung tissues from patients with IPF compared to control tissues (Fig. 1A), and that this 9 correlated (R = 0.89, $P = 7.9 \times 10^{-15}$) with an upregulated oxidative stress gene expression 10 signature (Fig. 1B), findings consistent with our previous observation of increased 11 pseudohypoxic HIF activity in IPF tissue (2).

We then investigated the HIF score calculated from the bronchoalveolar lavage (BAL) cell 12 13 transcriptome of 176 patients from 3 independent IPF cohorts: Freiburg (Germany), Siena 14 (Italy) and Leuven (Belgium) (8). Compared to healthy volunteers, the HIF score was 15 significantly increased in patients with IPF across all 3 cohorts (Fig. 1C; P < 0.001), whilst 16 there was no significant difference in the HIF score between smoker control and chronic 17 obstructive pulmonary disease (COPD) patients (Fig. 1D). The HIF score in BAL samples also correlated with an oxidative stress gene expression signature (R = 0.48, $P = 1.3 \times 10^{-11}$; Figure 18 1-figure supplement 1). Together these findings identify that a validated HIF score is increased 19 20 in both lung tissue and bronchoalveolar lavage cells from patients with IPF, and that this is 21 associated with increased oxidative stress.

Figure 1. A validated HIF score is increased across tissue compartments in patients with

2 **IPF.** (A) HIF GSVA scores from lung tissue from healthy controls and IPF patients 3 (GSE92592). Data are mean. *P < 0.05 by unpaired *t*-test. (B) The scatter plot for the

- 4 correlation between HIF score and oxidative stress score in IPF lung tissues (Pearson's
- correlation R = 0.89 and $P = 7.9 \times 10^{-15}$). (C) HIF GSVA scores from bronchoalveolar lavage 5
- 6 cells from 3 independent IPF cohorts (Freiburg, Siena and Leuven) compared to healthy donors
- 7 (GSE70867). Data are mean. ****P < 0.0001 by Dunnett's multiple comparisons test. (D) HIF
- 8 GSVA scores in BAL samples from patients with COPD compared to healthy smokers. Data
- 9 are mean. ns (not significant) by unpaired *t*-test, P > 0.05.
- 10

1



1 The HIF score in BAL predicts mortality in patients with IPF.

2 We next investigated whether the HIF score itself had prognostic value in this BAL cohort. 3 Classifying patients into score-high and score-low groups based on unsupervised hierarchical 4 clustering (Figure 2-figure supplement 1A; Fig. 2A), we identified that the HIF score 5 significantly predicted survival in the whole IPF cohort (Fig. 2B; hazard ratio, HR: 5.4; P =6 2.55×10^{-7}), with comparable trends in each individual cohort (Figure 2-figure supplement 1B-7 D). A high HIF score was a strong independent predictor (HR: 4.1, P < 0.001) of mortality 8 including in multivariate analysis with the physiological Gender, Age and Physiology (GAP) 9 score that uses commonly measured clinical and physiologic variables to predict mortality in 10 IPF (9) (Fig. 2C). Notably, the HIF score alone performed similarly to the GAP index (Fig. 11 2D), and when the HIF score and the GAP index were combined, the outcome prediction error 12 rate was reduced, indicating a significant added value (Fig. 2D).

13 To consider whether the change in HIF score might represent perturbations in cell types we 14 used annotated lung scRNA-seq data (10) as a reference to perform cell type deconvolution 15 analysis (CIBERSORTx) (11) of the bulk BAL cell gene profile. This identified macrophages (~70%) as the major cell type in BAL (Figure 2-figure supplement 1E), a proportion consistent 16 17 with the reported BAL macrophage differential count of this patient group (8). Comparing IPF 18 patients with high HIF and low HIF scores identified a significant reduction in the macrophage 19 proportion in HIF score high patients whilst the proportion of monocytes, B cells, dendritic 20 cells, endothelial cells and club cells were significantly increased (Fig. 2E).

Figure 2. The HIF score in BAL predicts mortality in patients with IPF. (A) Heat map of 1 2 the HIF score (ACOT7, ADM, ALDOA, CDKN3, ENO1, LDHA, MIF, MRPS17, NDRG1, 3 P4HA1, PGAM1, SLC2A1, TP11, TUBB6 and VEGFA) from bronchoalveolar lavage cells of 4 patients with IPF (GSE70867). Red indicates up-regulation and blue down-regulation. (B) 5 Kaplan-Meier plot showing the overall survival in IPF patients with low vs. high HIF scores in 6 BAL. Numbers below are n (%). P values, hazard ratio (HR), 95% confidence interval (CI) 7 and patient number (n) are indicated. (C) Multivariate analysis in IPF patients. HR (hazard 8 ratio), 95% CI (confidence interval), patient number (n) and P values are shown. (D) Prediction 9 error curves (indicating mean squared error in predicting survival status) were calculated for 10 the HIF score (blue) and Gender Age Physiology (GAP) index (green). Combining the HIF 11 score with GAP score (red) reduced prediction errors and resulted in better prediction. (E) Cell 12 deconvolution based on a signature matrix derived from the single-cell RNAseq found in 13 Reyfman et al was used to determine cell compositions in BAL samples from IPF patients with low vs. high HIF score (GSE70867). *P < 0.05 by unpaired *t*-test. **P < 0.01. 14





1 The HIF score is increased in peripheral blood mononuclear cells (PBMC) from IPF

2 patients.

Although BAL provides direct insight into molecular events in the lung milieu, whole blood and peripheral blood mononuclear cells (PBMC) have the advantage of being a less invasive sampling method (12) and so we next applied the HIF score to PBMC transcriptome datasets. We first applied the HIF gene expression signature (3) to a PBMC microarray dataset (GSE38958) (13, 14) encompassing 45 healthy controls and 70 IPF patients, identifying that the HIF score was increased in PBMC samples from IPF patients compared to health controls (Fig. 3A; P < 0.05).

10 We next calculated the HIF score in 2 PBMC microarray datasets of IPF patients with available longitudinal outcome data (GSE28042, Pittsburgh, n=75; and GSE27957, Chicago, 11 12 n=45) (15, 16). Classifying patients into score-high and score-low groups (Figure 3-figure 13 supplement 1A; Fig. 3B), we performed gene ontology (GO) and pathway enrichment analysis 14 of differentially expressed genes (DEGs), identifying enrichment for "reactive oxygen species 15 metabolic process" and "hallmark hypoxia" within the HIF score high group (Fig. 3C). 16 Furthermore, applying the HIF signature or oxidative stress gene expression signatures to this 17 dataset we identified a significant correlation, consistent with an increase in pseudohypoxic HIF activity in PBMCs (R = 0.48, $P = 3.5 \times 10^{-8}$; Fig. 3D). 18

To characterize the cell type composition within the PBMC RNA-seq data we next applied CIBERSORTx deconvolution using a leukocyte gene signature matrix developed by Newman and colleagues (17) as a reference (Figure 3-figure supplement 1B). Consistent with our observations in BAL samples, the proportion of monocytes identified in PBMCs was significantly increased in IPF patients with higher HIF scores (Fig. 3E; 44.4% *vs.* 37.1%; *P* < 0.0001) whilst the naïve CD4 T cell population and M2 Macrophages were decreased.

Figure 3. The HIF score is increased in PBMC from IPF patients. (A) HIF GSVA scores 1 2 in PBMC samples from healthy controls and IPF patients (GSE38958). Data are mean. *P <3 0.05 by unpaired *t*-test. (B) Heat map of the HIF score (ACOT7, ADM, ALDOA, CDKN3, ENO1, 4 LDHA, MIF, MRPS17, NDRG1, P4HA1, PGAM1, SLC2A1, TP11, TUBB6 and VEGFA) in the 5 microarray datasets GSE28042 and GSE27957. Red indicates up-regulation and blue down-6 regulation. (C) GO and pathway enrichment in HIF score high PBMC samples visualised on 7 a bar chart clustered by molecular functions, cellular component, biological process and 8 hallmark pathway. (D) The scatter plot for the correlation between HIF score and oxidative 9 stress score in PBMC samples (Pearson's correlation R = 0.48 and $P = 3.5 \times 10^{-8}$). (E) Cell 10 deconvolution based on a signature matrix from Newman et al was used to determine cell compositions in PBMC samples from IPF patients with low vs. high HIF score (GSE38958). 11 *P < 0.05. ***P < 0.001. ****P < 0.0001 by unpaired *t*-test. 12





1 The HIF score in PBMC from IPF patients predicts mortality.

- 2 Finally, we investigated whether the HIF score had prognostic value in PBMC. We identified 3 that a high HIF score predicted mortality in IPF patients from the Pittsburgh cohort (Figure 4-4 figure supplement 1A; HR: 13; P = 0.021), Chicago cohort (Figure 4-figure supplement 1B; 5 HR: 4.3; P = 0.047) and the combined cohort (Chicago+ Pittsburgh) (Fig. 4A; HR: 7.2; P =6 0.0073). A high HIF score in PBMC was also a strong independent predictor (HR: 5.9, P =7 0.003) of mortality in multivariate analysis with gender and age (Fig. 4B). Thus, HIF activity 8 assessed by a validated HIF metagene signature is a strong independent predictor of mortality 9 across tissue compartments in patients with IPF. 10 11
- 12 13
- 14



Figure 4. The HIF score in PBMC from IPF patients predicts mortality. (A) Kaplan-Meier 27 28 plots showing the overall survival in IPF patients with low vs. high HIF scores in PBMC from 29 the combined cohort (Chicago+ Pittsburg; GSE28042 and GSE27957). Numbers below are n 30 (%). P values, hazard ratio (HR), 95% confidence interval (CI) and patient number (n) are 31 indicated. (B) Multivariate analysis in IPF patients. HR (hazard ratio), 95% CI (confidence 32 interval), patient number (n) and P values are shown.

1 Discussion

We previously reported that altered collagen fibril nanoarchitecture is a core determinant of dysregulated ECM structure-function in human lung fibrosis (1), and that this dysregulation is promoted by pseudohypoxic HIF pathway activation (2). Here, we extend these observations to identify that HIF activity, as determined by a HIF gene signature, is increased in lung tissue, BAL, and PBMC samples from patients with IPF, that this strongly correlates with an oxidative stress signature consistent with pseudohypoxic HIF pathway activation, and that this is a strong independent predictor of mortality.

9 We selected a well-validated 15-gene expression HIF signature (3-6) to investigate HIF 10 activity. In multiple studies this 15-gene signature has been identified to strongly correlate with 11 independent hypoxia/HIF signatures (4-6). Furthermore, it was shown to be the best performer 12 in a comprehensive study assessing the robustness of different hypoxia/HIF signatures (18). 13 Importantly, it performed consistently across multiple pre-processing analytical pipelines (18), 14 so supporting potential utility as a biomarker of disease activity. Our identification of consistent 15 findings across tissue compartments in multiple patient cohorts is supportive of the need for 16 prospective clinical studies to investigate the clinical utility of the HIF signature as a biomarker 17 of disease activity. Although our study has focused upon patients with IPF, plausibly the 15-18 gene signature might have utility as a predictor of disease progression for any form of lung fibrosis. 19

20 Whilst hypoxia causes activation of the HIF pathway by inhibition of HIF prolyl 21 hydroxylases that hydroxylate HIF- α and target it for degradation via the ubiquitin proteasome 22 system, our recent study in IPF identified evidence of increased HIF activity under normoxic 23 conditions (pseudohypoxic HIF activation) as a consequence of oxidative stress. Although we 24 cannot exclude the possibility that hypoxia status contributed to HIF score in the cohorts 25 studied, the HIF score was a strong independent predictor of mortality in multivariate analysis

with the physiological Gender, Age and Physiology (GAP) score, concordant with HIF pathway activation status being an independent biomarker of disease activity rather than simply a reflection of hypoxia status determined by disease severity. Additionally, we identified that the HIF score correlated with an oxidative stress gene expression signature across multiple tissue compartments, consistent with systemic oxidative stress modulating HIF pathway activity status (19).

7 Oxidative stress has been strongly implicated as an important profibrotic mechanism in the 8 lungs, with N-acetylcysteine (NAC), a precursor of the anti-oxidant glutathione, proposed as a 9 potential treatment for IPF. However, the role of NAC for IPF treatment remains controversial, 10 with the PANTHER study identifying no benefit compared to placebo (20). Whilst NAC is not 11 recommended as a treatment for IPF, meta-analyses continue to propose potential clinical 12 utility (21). Further consideration is warranted into whether patients with a high HIF-score 13 might benefit from NAC treatment. In support of the potential for stratified NAC treatment 14 guided by HIF-score, post-hoc exploratory analysis of patients from the PANTHER study has 15 proposed that in individuals with a polymorphism of the TOLLIP gene (TT rs3750920), NAC 16 may be efficacious by decreasing TLR-4 oxidant dependent signaling (22).

17 We identified an increased monocyte proportions in BAL and PBMC samples with a high 18 HIF score. A number of recent reports have proposed peripheral blood monocyte count as a 19 predictor of IPF disease behaviour (23-25), with elevated monocyte count associated with 20 increased risks of IPF progression, hospitalization, and mortality. Further investigation is 21 required to confirm whether an increased HIF score reflects a dysregulated monocyte 22 population and/or increased monocyte population. HIF activity is reported to regulate the 23 development and function of hematopoietic progenitor/stem cells, with HIF pathway activation 24 promoting metabolic and phenotypic reprogramming of myeloid cells (26). Consistent with 25 these reports using a clinically relevant chronic intermittent hypoxia (CIH) model, Alvarez-

1 Martins and colleagues were able to show an increase in monocytes upon CIH exposure (27), 2 potentially via modulating monocyte differentiation, proliferation, and survival (28, 29). 3 Furthermore, it has been recently reported (30) that monocytes from IPF patients are 4 phenotypically distinct compared to age matched controls and this was associated with an 5 increase in the level of macrophage colony-stimulating factor 1 (CSF-1) in serum. Notably, 6 CSF-1 gene transcription has previously been reported (31) to be dependent upon HIF 7 transcriptional activity.

8 In summary, we identify that a validated HIF score is increased across tissue compartments 9 in patients with IPF, that this strongly correlates with oxidative stress consistent with 10 pseudohypoxic HIF activation, and that an increase in this HIF score is a strong independent 11 predictor of mortality. Prospective validation of these findings is warranted which could inform 12 stratified approaches for the treatment of lung fibrosis.

Methods 1

2 Data preparation, analysis for Gene Ontology (GO) and pathway enrichment

3 Microarray data of BAL and PBMC samples are publicly available in NCBI in the Gene 4 Expression Omnibus (GEO) with the accessions GSE70867 (8), GSE73395 (8), GSE38958 5 (13, 14), GSE28042 and GSE27957 (15, 16). Raw microarray files were downloaded and 6 imported into Rstudio using the GEOquery (v2.56.0). The normalization of raw data depended 7 platform: Agilent microarrays were on the analytical normalized using the 8 normalizeBetweenArrays function in limma (v3.44.3) and Affymetrix data were normalized 9 using the rma function in affy (v1.66.0) packages. Microarray probe IDs were mapped to Gene 10 symbol according to the GPL annotation files provided in NCBI. Probes mapped to multiple 11 gene symbols were removed and genes mapped to multiple probe IDs were summarized by 12 calculating the mean. The ComBat function in sva (v3.36.0) was then used to correct the 13 technical batch effect. Differential expression analysis was performed using the limma package 14 (version 3.40.6). Genes with adjusted P value (Benjamini-Hochberg) less than 0.05 were 15 considered as differentially expressed genes (DEGs). RNA-seq data of control vs. IPF lung 16 tissues were downloaded from GEO with the accession GSE92592 (32). Raw read counts were 17 normalized and analyzed by using R package of DESeq2. Pathway enrichment analysis was 18 performed in Metascape with default parameters. Adjusted P value < 0.05 as the cut-off 19 criterion was considered statistically significant.

20

21 Oxidative stress and HIF score calculation, and grouping

22 An oxidative stress signature was established previously based on upregulated genes in IPF 23 cell populations from HALLMARK REACTIVE OXYGEN SPECIES PATHWAY gene 24 set (ABCC1, CDKN2D, FES, GCLC, GCLM, GLRX2, HHEX, IPCEF1, JUNB, LAMTOR5, LSP1, MBP, MGST1, MPO, NDUFA6, PFKP, PRDX1, PRDX2, PRDX4, PRNP, SBNO2, 25

1 SCAF4, SOD1, SOD2, RXN1, TXN, TXNRD1) (2). Buffa and colleagues summarized a HIF 2 signature consisting of 15 genes (ACOT7, ADM, ALDOA, CDKN3, ENO1, LDHA, MIF, 3 MRPS17, NDRG1, P4HA1, PGAM1, SLC2A1, TPI1, TUBB6 and VEGFA) by combining gene 4 function analysis and gene co-expression analysis in vivo (3). Based on these signatures, the 5 oxidative stress score and HIF score for each sample were calculated by using gene set variation 6 analysis in the GSVA (v1.36.2) package (7). To classify HIF scores, we employed 7 unsupervised hierarchical clustering (ward.D) to cluster samples according to the HIF scores. 8 Student's *t*-test was used to evaluate the statistical difference between high *vs.* low score groups.

9

10 Survival analysis

Univariate and multivariate Cox proportional hazard models were used to assess the hazard ratio (HR) of each parameter via the survminer (v0.4.8). We performed log-rank tests to compare Kaplan-Meier survival curves between groups with high or low scores by survival (v3.2-3). Prediction error curves of each prognostic model were generated from pec (v2019.11.03).

16

17 Single-cell data analysis

The single-cell dataset was downloaded from GEO (Accession: GSE122960). Seurat (v2.2)
was used to filter cells, cluster and perform t-distributed stochastic neighbor embedding (tSNE) reduction and identify cell types by marker genes described by Reyfman and colleagues
(10).

22

23 **CIBERSORT** analysis

We used CIBERSORTx to estimate the percentage of 22 types of infiltrating immune cells in
each PBMC samples with default parameters (11). For BAL samples, the first step was to create

the signature matrix of each cell type in IPF patients from single-cell gene expression profiling.
The imputation of bulk BAL sample used the signature matrix with S-batch correction and
1000 permutations for significance analysis. Student's *t*-test was used to evaluate the statistical
difference in each cell population between the two conditions. *P*-values were adjusted for
multiple testing using the Benjamini-Hochberg method.

6

7 Statistics

8 Statistical analyses were performed in GraphPad Prism v7.02 (GraphPad Software Inc., San 9 Diego, CA) unless otherwise indicated. We evaluated the correlations between oxidative stress 10 score and HIF score using Pearson's correlation. Normality of distribution was assessed using 11 the D'Agostino-Pearson normality test. Statistical analyses of single comparisons of two 12 groups utilised Student's t-test or Mann-Whitney U-test for parametric and non-parametric data, 13 respectively. Where appropriate, individual Student's *t*-test results were corrected for multiple 14 comparisons using the Holm-Sidak method. Results were considered significant if P < 0.05, 15 where *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001.

16

17 Code availability

18 Codes were implemented in R and have been deposited in GitHub:
19 https://github.com/yz3n18/HIF_IPF.

1 Acknowledgements

- 2 This project was supported by Medical Research Council [MR/S025480/1], an Academy of
- 3 Medical Sciences/the Wellcome Trust Springboard Award [SBF002\1038] and the AAIR
- 4 Charity. YZ was supported by an Institute for Life Sciences PhD Studentship.

5

6 **Conflict of interest**

7 The authors declare that they have no relevant conflict of interest.

Supplementary Figures 1

Figure 1-figure supplement 1. A validated HIF score is increased across tissue 2 3 compartments in patients with IPF. The scatter plot for the correlation between HIF score 4 and oxidative stress score calculated using GSVA in IPF BAL samples (GSE70867) (Pearson's correlation R = 0.48 and $P = 1.3 \times 10^{-11}$).

- 6
- 7



Figure 2-figure supplement 1. The HIF score in BAL predicts mortality in patients with 1

2 **IPF.** (A) Hierarchical clustering dendrogram plot of BAL samples from IPF patients 3 (GSE70867) according to HIF scores. (B-D) Kaplan-Meier plot showing the overall survival

4 in IPF patients of each cohort with low vs. high HIF scores in BAL. Numbers below are n (%).

5 P values, hazard ratio (HR), 95% confidence interval (CI) and patient number (n) are indicated.

- 6 (E) Violin plot showing cell deconvolution proportions based on a signature matrix derived
- 7 from the single-cell RNAseq found in Reyfman et al in BAL samples from IPF patients with
- 8 (GSE70867)
- 9
- 10



1 Figure 3-figure supplement 1. The HIF score is increased in PBMC from IPF patients. (A)

2 Hierarchical clustering dendrogram plot of PBMC samples (GSE28221) from IPF patients 3 according to HIF scores. (B) Violin plot showing cell compositions of PBMC using 4 CIBERSORTx analysis.



1 Figure 4-figure supplement 1. The HIF score in PBMC from IPF patients predicts

2 mortality. Kaplan-Meier plot showing the overall survival in IPF patients of in (A) Pittsburgh

3 and (B) Chicago cohorts with low vs. high HIF scores in PBMC (GSE28042 and GSE27957). 4 Numbers below are n (%). P values, hazard ratio (HR), 95% confidence interval (CI) and

- 5 patient number (n) are indicated.
- 6 7



References

2	1. Jones MG, Andriotis OG, Roberts JJ, Lunn K, Tear VJ, Cao L, Ask K, Smart DE, Bonfanti
3	A, Johnson P, Alzetani A, Conforti F, Doherty R, Lai CY, Johnson B, Bourdakos KN,
4	Fletcher SV, Marshall BG, Jogai S, Brereton CJ, Chee SJ, Ottensmeier CH, Sime P,
5	Gauldie J, Kolb M, Mahajan S, Fabre A, Bhaskar A, Jarolimek W, Richeldi L,
6	O'Reilly KM, Monk PD, Thurner PJ, Davies DE. Nanoscale dysregulation of collagen
7	structure-function disrupts mechano-homeostasis and mediates pulmonary fibrosis.
8	<i>Elife</i> 2018; 7: e36354.
9	2. Brereton C, Yao L, Zhou Y, Vukmirovic M, Bell J, Ridley RA, Davies ER, Dean LSN,
10	Andriotis OG, Conforti F, Mohammed S, Wallis T, Tavassoli A, Ewing R, Alzetani
11	A, Marshall BG, Fletcher SV, Thurner PJ, Fabre A, Kaminski N, Richeldi L, Bhaskar
12	A, Loxham M, Davies DE, Wang Y, Jones MG. Pseudohypoxic HIF pathway
13	activation dysregulates collagen structure-function in human lung fibrosis. <i>eLife</i> 2022
14	Accepted/in press
15	3. Buffa FM, Harris AL, West CM, Miller CJ. Large meta-analysis of multiple cancers
16	reveals a common, compact and highly prognostic hypoxia metagene. Br J Cancer
I7	2010; 102: 428-435.
18	4. Ye Y, Hu Q, Chen H, Liang K, Yuan Y, Xiang Y, Ruan H, Zhang Z, Song A, Zhang H,
19	Liu L, Diao L, Lou Y, Zhou B, Wang L, Zhou S, Gao J, Jonasch E, Lin SH, Xia Y,
20	Lin C, Y ang L, Mills GB, Liang H, Han L. Characterization of Hypoxia-associated
21	Molecular Features to Aid Hypoxia-Targeted Therapy. Nat Metab 2019; 1: 431-444.
22	J. Bhandari V, Hoey C, Liu LY, Lalonde E, Kay J, Livingstone J, Lesuri K, Shian YJ, Vujcic T. Huang V. Egnirity SMG, Heisler J E, Vousif F, Huang V, Vomegushi TN, Vos
25 24	CO Sabelnykova VV Fraser M Chua MI K van der Kwast T Liu SK Boutros PC
2 1 25	Bristow RG Molecular landmarks of tumor hypoxia across cancer types. Nat Ganat
25	2019: 51: 308-318
27	6 Bhandari V Li CH Bristow RG Boutros PC Consortium P Divergent mutational
28	processes distinguish hypoxic and normoxic tumours. <i>Nat Commun</i> 2020: 11: 737.
29	7. Hanzelmann S. Castelo R. Guinney J. GSVA: gene set variation analysis for microarray
30	and RNA-seq data. BMC Bioinformatics 2013; 14: 7.
31	8. Prasse A, Binder H, Schupp JC, Kayser G, Bargagli E, Jaeger B, Hess M, Rittinghausen S,
32	Vuga L, Lynn H, Violette S, Jung B, Quast K, Vanaudenaerde B, Xu Y, Hohlfeld JM,
33	Krug N, Herazo-Maya JD, Rottoli P, Wuyts WA, Kaminski N. BAL Cell Gene
34	Expression Is Indicative of Outcome and Airway Basal Cell Involvement in
35	Idiopathic Pulmonary Fibrosis. American journal of respiratory and critical care
36	<i>medicine</i> 2019; 199: 622-630.
37	9. Ley B, Ryerson CJ, Vittinghoff E, Ryu JH, Tomassetti S, Lee JS, Poletti V, Buccioli M,
38	Elicker BM, Jones KD, King TE, Jr., Collard HR. A multidimensional index and
39	staging system for idiopathic pulmonary fibrosis. Annals of internal medicine 2012;
40	156: 684-691.
41	10. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S,
42	Fernandez R, Akbarpour M, Chen CI, Ren Z, Verma R, Abdala-Valencia H, Nam K,
43	Chi M, Han S, Gonzalez-Gonzalez FJ, Soberanes S, Watanabe S, Williams KJN,
44 45	FIOZAK AS, NICHOISON II, MOrgan VK, WINTER DK, HINCHCHITT M, HRUSCH CL, GUZY
45 46	KD, Bonnam CA, Sperling AI, Bag K, Hamanaka KB, Mutlu GM, Yeldandi AV,
40 47	Warshall SA, Shilaulard A, Amaral LAN, Periman H, Sznajder JI, Argento AC,
4/ 10	SM Cottordi CL Dudinger CDS Migherin AV Single Coll Transprintensis Analysis
40	Sivi, Gottardi CJ, Budinger GKS, ivirsitarin AV. Single-Cen Transcriptomic Analysis

1	of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. Am J
2	Respir Crit Care Mea 2019; 199: 151/-1550.
3 4	MS. Esfahani MS. Luca BA. Steiner D. Diehn M. Alizadeh AA. Determining cell
5	type abundance and expression from bulk tissues with digital cytometry. <i>Nat</i>
6	Biotechnol 2019: 37: 773-782.
7	12. Maher TM. Blood-based Diagnosis of Idiopathic Pulmonary Fibrosis. Fantasy or Reality?
8	American journal of respiratory and critical care medicine 2016; 194: 1182-1184.
9	13. Huang LS, Berdyshev EV, Tran JT, Xie L, Chen J, Ebenezer DL, Mathew B, Gorshkova
10	I, Zhang W, Reddy SP, Harijith A, Wang G, Feghali-Bostwick C, Noth I, Ma SF,
11	Zhou T, Ma W, Garcia JG, Natarajan V. Sphingosine-1-phosphate lyase is an
12	endogenous suppressor of pulmonary fibrosis: role of S1P signalling and autophagy.
13	<i>Thorax</i> 2015; 70: 1138-1148.
14	14. Huang LS, Mathew B, Li H, Zhao Y, Ma SF, Noth I, Reddy SP, Harijith A, Usatyuk PV,
15	Berdyshev EV, Kaminski N, Zhou T, Zhang W, Zhang Y, Rehman J, Kotha SR,
10	Gurney IO, Parinandi NL, Lussier YA, Garcia JG, Natarajan V. The mitochondrial
l / 10	cardiolipin remodeling enzyme hysocardiolipin acyltransferase is a novel target in
10	pullionary horosis. American journal of respiratory and critical cure medicine 2014, 180, 1402–1415
20	15 Herazo-Maya ID Noth I Duncan SR Kim S Ma SF Tseng GC Feingold F Juan-
20	Guardela BM, Richards TL Lussier Y, Huang Y, Vii R, Lindell KO, Xue L Gibson
22	KF. Shapiro SD. Garcia JG. Kaminski N. Peripheral blood mononuclear cell gene
23	expression profiles predict poor outcome in idiopathic pulmonary fibrosis. Science
24	translational medicine 2013; 5: 205ra136.
25	16. Huang Y, Ma SF, Vij R, Oldham JM, Herazo-Maya J, Broderick SM, Strek ME, White
26	SR, Hogarth DK, Sandbo NK, Lussier YA, Gibson KF, Kaminski N, Garcia JG, Noth
27	I. A functional genomic model for predicting prognosis in idiopathic pulmonary
28	fibrosis. BMC pulmonary medicine 2015; 15: 147.
29	17. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M,
30	Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles.
31	Nature methods 2015; 12: 453-457.
32 22	18. Fox NS, Starmans MH, Halder S, Lambin P, Boutros PC. Ensemble analyses improve
33 34	Righter Biological Solution State St
35	19 Masson N Singleton RS Sekirnik R Trudgian DC Ambrose I I Miranda MX Tian
36	YM. Kessler BM. Schofield CL Ratcliffe PL The FIH hydroxylase is a cellular
37	peroxide sensor that modulates HIF transcriptional activity. <i>EMBO Rep</i> 2012: 13:
38	251-257.
39	20. Idiopathic Pulmonary Fibrosis Clinical Research N, Martinez FJ, de Andrade JA,
40	Anstrom KJ, King TE, Jr., Raghu G. Randomized trial of acetylcysteine in idiopathic
41	pulmonary fibrosis. N Engl J Med 2014; 370: 2093-2101.
42	21. Feng F, Zhang J, Wang Z, Wu Q, Zhou X. Efficacy and safety of N-acetylcysteine
43	therapy for idiopathic pulmonary fibrosis: An updated systematic review and meta-
44	analysis. <i>Exp Ther Med</i> 2019; 18: 802-816.
45	22. Oldham JM, Ma SF, Martinez FJ, Anstrom KJ, Raghu G, Schwartz DA, Valenzi E, Witt
46	L, Lee C, Vij K, Huang Y, Strek ME, Noth I, Investigators IP. TOLLIP, MUC5B, and
4/ 10	the Response to N-Acetylcysteine among Individuals with Idiopathic Pulmonary
4ð 40	FIDIOSIS. Am J Kespir Crit Care Mea 2015; 192: 14/5-1482.
47 50	Aren K Sun I Koloms K Lee I Baral I Kronski I Zhao H Herzog F Martinez FI
50	Aren K, Sun J, Koloms K, Lee J, Darar J, Kropski J, Zhao H, Herzog E, Martinez FJ,

1 2	Moore BB, Hinchcliff M, Denny J, Kaminski N, Herazo-Maya JD, Shah NH, Khatri P. Increased monocyte count as a cellular biomarker for poor outcomes in fibrotic
3 4	diseases: a retrospective, multicentre cohort study. <i>The Lancet Respiratory medicine</i> 2010: 7: 407-508
- - -5	2019, 7. 497-500. 24 Karampiteakos T. Torrisi S. Antoniou K. Manali F. Karbila I. Panajaannau O.
5	24. Karampiisakos I, Tomisi S, Antoniou K, Manan E, Korona I, Lapatoannou O, Samnsonas E Katsaras M Vasarmidi E Panakosta D Domuri K Eouka E Organtzis
07	L Daniil Z Dimeas I Kirgou P. Gourgoulianis KI Pananikalaou IC Markanoulou K
8	Kounti G. Tsanakidou E. Panadonoulou F. Tatsis K. Gogali A. Kostikas K. Tzilas V.
0	Chrysikos S. Paniris S. Bouros D. Kreuter M. Tzouvelekis A. Increased monocyte
9 10	count and red cell distribution width as prognostic biomarkers in patients with
11	Idionathic Pulmonary Fibrosis <i>Resnir Res</i> 2021: 22: 140
12	25 Kreuter M Lee IS Tzouvelekis A Oldham IM Molyneaux PL Weycker D Atwood M
13	Kirchgaessler KU Maher TM Monocyte Count as a Prognostic Biomarker in Patients
14	with Idionathic Pulmonary Fibrosis Am J Respir Crit Care Med 2021: 204: 74-81
15	26. Semba H. Takeda N. Isagawa T. Sugiura Y. Honda K. Wake M. Miyazawa H.
16	Yamaguchi Y. Miura M. Jenkins DM. Choi H. Kim JW. Asagiri M. Cowburn AS.
17	Abe H, Soma K, Kovama K, Katoh M, Savama K, Goda N, Johnson RS, Manabe I.
18	Nagai R, Komuro I. HIF-1alpha-PDK1 axis-induced active glycolysis plays an
19	essential role in macrophage migratory capacity. Nat Commun 2016; 7: 11635.
20	27. Alvarez-Martins I, Remedio L, Matias I, Diogo LN, Monteiro EC, Dias S. The impact of
21	chronic intermittent hypoxia on hematopoiesis and the bone marrow
22	microenvironment. Pflugers Archiv : European journal of physiology 2016; 468: 919-
23	932.
24	28. Rathinam C, Poueymirou WT, Rojas J, Murphy AJ, Valenzuela DM, Yancopoulos GD,
25	Rongvaux A, Eynon EE, Manz MG, Flavell RA. Efficient differentiation and function
26	of human macrophages in humanized CSF-1 mice. <i>Blood</i> 2011; 118: 3119-3128.
27	29. Roiniotis J, Dinh H, Masendycz P, Turner A, Elsegood CL, Scholz GM, Hamilton JA.
28	Hypoxia prolongs monocyte/macrophage survival and enhanced glycolysis is
29	associated with their maturation under aerobic conditions. Journal of immunology
30	2009; 182: 7974-7981.
31	30. Fraser E, Denney L, Antanaviciute A, Blirando K, Vuppusetty C, Zheng Y, Repapi E,
32	Iotchkova V, Taylor S, Ashley N, St Noble V, Benamore R, Hoyles R, Clelland C,
33	Rastrick JMD, Hardman CS, Alham NK, Rigby RE, Simmons A, Rehwinkel J, Ho
34	LP. Multi-Modal Characterization of Monocytes in Idiopathic Pulmonary Fibrosis
35	Reveals a Primed Type I Interferon Immune Phenotype. Front Immunol 2021; 12:
36	623430.
3/	31. Chaturvedi P, Gilkes DM, Takano N, Semenza GL. Hypoxia-inducible factor-dependent
38 20	signaling between triple-negative breast cancer cells and mesenchymal stem cells
39 40	promotes macrophage recruitment. Proc Natl Acad Sci U S A 2014; 111: E2120-2129.
+0 ∕11	52. Scharer IviJ, White TA, Hjillia K, Haak AJ, Ligresu G, Alkinson EJ, Oberg AL, Birch J, Salmonowicz H. Zhu V. Mazula DI, Brooks DW, Euhrmann Straissnicz H
+1 ∕\?	Dirtschalava T. Prakash VS. Tehkonia T. Dobbins DD. Aubry MC. Dassos IF
ד∠ ⊿3	Kirkland II Tschumperlin DI Kita H LeBrasseur NK Cellular senescence mediates
44 44	fibrotic pulmonary disease Nat Commun 2017. 8. 14532
45	norone pullionary discuse. Ital Comman 2017, 0. 17552.