EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Research letter

Circulating RNA differences between patients with stable and progressive IPF

Britt Clynick, Helen E. Jo, Tamera J. Corte, Ian N. Glaspole, Christopher Grainge, Peter M.A. Hopkins, Paul N. Reynolds, Sally Chapman, E. Haydn Walters, Christopher Zappala, Gregory J. Keir, Wendy A. Cooper, Annabelle M. Mahar, Samantha Ellis, Nicole S. Goh, Svetlana Baltic, Marisa Ryan, Dino B.A. Tan, Yuben P. Moodley

Please cite this article as: Clynick B, Jo HE, Corte TJ, *et al*. Circulating RNA differences between patients with stable and progressive IPF. *Eur Respir J* 2020; in press (https://doi.org/10.1183/13993003.02058-2019).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

Title: Circulating RNA differences between patients with stable and progressive IPF

Authors: Britt Clynick[1,2], Helen E Jo[3,4], Tamera J Corte[3,4], Ian N Glaspole[5,6], Christopher Grainge[7,8], Peter MA Hopkins[9,10], Paul N Reynolds[11,12], Sally Chapman[12], E. Haydn Walters[6,13,14,15], Christopher Zappala[9], Gregory J Keir[9], Wendy A Cooper[3,4,16], Annabelle M Mahar[3,4], Samantha Ellis[6], Nicole S Goh[17,18], Svetlana Baltic[1,2], Marisa Ryan[1,2], Dino BA Tan[1,2], Yuben P Moodley[1, 2,19]

Institutional Addresses:

1. Institute for Respiratory Health, Nedlands, Western Australia.
2. University of Western Australia, Crawley, Western Australia.
3. University of Sydney, Camperdown, New South Wales.
4. Royal Prince Alfred Hospital, Camperdown, New South Wales.
5. Monash University, Clayton, Victoria.
6. Alfred Hospital, Melbourne, Victoria.
7. University of Newcastle, Callaghan, New South Wales.
8. John Hunter Hospital, New Lambton Heights, New South Wales.
9. University of Queensland, St Lucia, Queensland.
10. Prince Charles Hospital, Chermside, Queensland.
11. University of Adelaide, Adelaide, South Australia.
12. Royal Adelaide Hospital, Adelaide, South Australia.
13. University of Tasmania, Hobart, Tasmania.
14. University of Melbourne, Parkville, Victoria.
15. Royal Hobart Hospital, Hobart, Tasmania.
16. Western Sydney University, Sydney, New South Wales.

17. Austin Hospital, Heidelberg, Victoria.

18. Institute of Breathing and Sleep, Heidelberg, Victoria.

19. Fiona Stanley Hospital, Murdoch, Western Australia.

Corresponding author:

Yuben P Moodley

School of Medicine, University of Western Australia

Level 2, Harry Perkins Institute of Medical Research,

Fiona Stanley Hospital Campus,

5 Robin Warren Drive, Murdoch WA 6150

Email: yuben.moodley@uwa.edu.au

Phone: +61 414 383 338

Summary

We explored gene expression profile differences found in the circulation of IPF patients versus healthy controls, identifying 7 potentially relevant transcripts, of which 5 were expressed in higher concentrations in progressive versus stable IPF, potentially providing insight into disease pathogenesis and progression.

Research Letter

Idiopathic pulmonary fibrosis (IPF) is a chronic disease characterised by progressive decline in pulmonary function. The rate of decline can vary, with some patients remaining stable over longer periods of time and others rapidly progressing.[1]tThe variable progression of this disease makes it difficult to elucidate pathogenic pathways involved in the initiation and progression of IPF. Advances in high-throughput gene-expression analyses have led to improvements in our understanding of disease biology and prognostic gene signatures. We

hypothesise that IPF has a unique circulatory transcriptional profile compared to healthy controls, with additional differences between stable and progressive disease likely related to disease pathogenesis.

The study cohort consisted of consented patients from the Australian IPF Registry with clinical/physiological/radiographical findings consistent with the diagnosis of IPF. All work was approved by the Royal Perth Hospital Ethics Committee (HREC/2011-138), and the Sydney Local Health Network (HREC/15/RPAH/28). Baseline FVC and DLco were assessed ±6months from the time of blood collection, and the longitudinal FVC and DLco trajectories were determined ±6-12months from the baseline lung function using a linear regression model. A decline in FVC≥10% and/or DLco≥15% within 6-12months of baseline was used to define progressive IPF. No patients were on anti-fibrotic medications at blood collection.

An initial 10 patients from each group had plasma isolated and RNA extracted. Expression of over 135,000 transcripts were analysed by microarray (Human Clariom D; ThermoFisher Scientific), and expression profiles were compared between stable and progressive IPF samples. The top targets with a minimum 2-fold difference between the two IPF groups were identified and droplet digital PCR (ddPCR; BioRad) was used to validate expression differences and compare absolute expression measurements between an independent cohort of stable (n=33) and progressive (n=24) IPF patients and disease-free healthy controls (n=15). In contrast to other methods, ddPCR provides an absolute, objective quantification of the number of mRNA transcripts with high precision. This is based on partitioning samples into thousands of uniformly nanolitre-sized droplets, undergoing end-point PCR, and template concentration was determined using Poisson's statistical analysis of the ratio of positive (containing amplified target) to negative (no detectable amplified target) droplets detectable.

Five independent formalin-fixed paraffin-embedded (FFPE) IPF, 4 healthy lung control FFPE specimens, IPF and normal fibroblast cells lines as well as 5 COPD plasma samples used as a disease control group, were all analysed to confirm expression of genes detected in IPF patient circulation. A549 respiratory adenocarcinoma cell line was included as a positive control for gene expression analyses and for assay quality control. P-values for relative gene expression levels of each transcript detected by the microarray and ddPCR were generated using one-way analysis of variance (ANOVA) adjusted for multiple comparisons (Kruskal-Wallis with Dunn's multiple comparisons test) or Mann-Whitney test. The predictive performance of gene expression levels was examined using Cox proportional hazards regression analysis adjusting for age FVC baseline, gender and GAP stage. Global inter- and intra-group variability of the data was carried out by performing a principal component analysis (PCA). Statistical analyses were performed on SPSS version 24.

The mean age was 71±7yr in the IPF stable group (n=33; 21 males); 65±10yr in the IPF progressive group (n=24; 15 males); and 62±10yr in the healthy control group (n=15; 8 males). Lung function at baseline in the stable group was FVC 79±26% predicted and DLco 49±15% predicted, versus FVC of 78±18% predicted and DLco 43±13% predicted in the progressive group. There were 15 never smokers (stable=12; progressive=3), 38 ex-smokers (stable=19; progressive=19), 3 current smokers (stable=2; progressive=1) and 1 unknown.

From a total of 135,750 transcripts analysed in the microarray, 127 genes were differentially expressed between stable and progressive IPF patients. The microarray data was further filtered, and the most abundant 8 transcripts with >2-fold gene expression difference between IPF groups were selected for validation.

Validation by ddPCR confirmed 7 of the 8 transcripts (*TAF2*, *NT5C2*, *JAK1*, *TAOK1*, *TRAM1*, *RP11-726G23.6* and *MIR6841*) were differentially expressed between IPF and

healthy controls, of which 5 of the transcripts (*TAF2*, *NT5C2*, *JAK1*, *TRAM1*, and *RP11-726G23.6*) were observed at higher concentrations in progressive versus stable IPF samples (Figure 1A), with strongest evidence for *TAF2* (p=0.0413). ddPCR verification also confirmed higher expression of the 7 transcripts in IPF lung tissue and IPF fibroblasts relative to healthy lung tissue and fibroblasts derived from normal controls. Immunolocalisation staining by immunohistochemistry (IHC) was carried out on the 5 IPF lung FFPE samples to characterise significantly expressed *TAF2*. Stronger TAF2 expression was observed in the cytoplasm of bronchial epithelial cells, alveolar epithelial cells, smooth muscle cells and fibroblasts in IPF tissue (Figure 1B) relative to healthy lung (Figure 1C). *TAF2* expression was predictive of increased mortality ($p<0.05$) on multivariate Cox regression. PCA revealed that *TAF2* and *RP11-726G23.6* expression had a positive predictive relationship with IPF progression status (p=0.036).

This study aimed to investigate the circulating transcriptome in stable versus progressive IPF. Analysis of gene expression identified 7 transcripts in the plasma (confirmed in IPF lung tissue) that were differentially expressed in IPF compared to healthy controls, with a trend of increased circulating levels of these transcripts in progressive versus stable IPF. Specifically, expression of *TAF2* was significantly higher in progressive versus stable IPF, which may represent a marker indicative of disease progression. Interestingly, gene expression analysis in circulation of a cohort of COPD patients showed significantly increased expression relative to healthy controls in all but *MIR6841* where the strength of evidence was more moderate (p=0.055), further indicating the likely relevance of these genes in a chronic fibrosing and remodelling lung diseased setting.

TAF2 (TATA-Box Binding Protein Associated Factor-2) encodes an integral component of the core transcriptional machinery for RNA polymerase II. TAF proteins regulate differentiation and proliferation, important factors in the pathogenesis of IPF.[2] Interestingly,

lung data from the Human Protein Atlas Tissue Gene Expression Profiles dataset has reported the expression of *TAF2* predominantly in pneumocytes and endothelial cells making up 50-75% of the expression relative to all cell types in the lung.[3]

NT5C2 (5'-Nucleotidase, Cytosolic-II) encodes a hydrolase that serves an important role in cellular purine metabolism, and cell survival.[4] The RNAseq study by Nance *et al.* (2014)[5] reported higher expression of the *NT5C2* gene in IPF lung tissue relative to healthy controls. Its function in the maintenance of intracellular nucleotide pool homeostasis has been described in neurological disorders and leukaemias and requires further investigation in IPF.[6]

JAK1 (Janus Kinase 1), is a tyrosine kinase protein involved in the activation of several signal transduction pathways involved in differentiation, proliferation, survival and migration. STAT3 acting down-stream of JAK1 is a key regulator of fibroblast phenotype.[7]

TAOK1 (Thousand And One Amino Acid Protein Kinase-1) encodes a protein kinase involved in the stress-activated MAPK pathway, regulating a DNA damage response and apoptosis.[8, 9] The MAPK signalling cascade is known to regulate cellular process involved in fibrogenesis such as EMT. Although *TAOK1* has never been described in IPF, it has been reported to exacerbate liver fibrosis via the overexpression of α-smooth muscle actin (α-SMA).[10]

TRAM1 (Translocation Associated Membrane Protein-1), encodes a protein which forms part of the mammalian endoplasmic reticulum (ER), facilitating the translocation of proteins across its membrane. TRAM1 is upregulated under conditions of ER stress which might be relevant in IPF.[11, 12]

RP11-726G23.6 and *MIR6841* are non-coding genes that have lost their ability to code for proteins. *MIR6841* specifically is a non-coding microRNA (miRNA) likely involved in post-

transcriptional regulation of gene expression.[13] Although not described in IPF, *MIR6841* is associated with *RICTOR* (RPTOR Independent Companion of MTOR Complex-2), a protein coding gene that forms a subunit of mTORC2 (Mammalian Target of Rapamycin Complex-2) which is known to be associated with pulmonary fibrosis.[14]

Taken together, the above-mentioned circulatory genes were upregulated in IPF relative to healthy controls, with significantly higher expression of *TAF2* in progressive versus stable IPF. Notably, this was supported by the RNAseq study by Nance *et al.* (2014)[5], reporting higher expression of *TAF2*, *NT5C2*, and *TRAM1* genes in IPF versus healthy lung tissue in a smaller cohort (n=8 IPF patients vs n=7 healthy controls) with undefined disease progression status.

Another relevant study explored the predictive outcome of circulating RNA in IPF using a 52-gene expression microarray.[15] Herazo-Maya and colleagues revealed the significant improvement in outcome predictive accuracy when their 52-gene risk profiles were added to a patients GAP index. Unlike our study, their measured outcomes included transplant-free survival and mortality, using pooled data relating to age, gender, percent predicted forced vital capacity (FVC%) and immunosuppressive therapy, and not absolute percentage decline in lung function test as per our study. Furthermore, their study based the gene risk profiles on a 52-gene signature in RNA isolated from PBMC, whereas our study only explored genes free in the circulation. This could be a possible explanation as to why neither studies identified overlapping genes. Another strength of our study was the inclusion of a second chronic lung disease, COPD, which interestingly displayed similarly increased levels of these circulatory genes. The ability to characterise circulatory biomarkers could increase our understanding not only of the pathogenesis and progression of IPF, but potentially other chronic fibrosing/remodelling lung diseases such as COPD. Validation of our findings in an independent cohort will need to be carried out, and although one statistically significant

difference was observed in gene expression between progressive and stable IPF, trends were observed, and larger cohorts will be needed to confirm true significance. Most of the implicated proteins have potential to be involved in pathogenic mechanisms of IPF, so future functional studies focused on these biomarkers are warranted.

Acknowledgements

The study is supported by NHMRC grants (APP 1147776 and APP 1066128) and the Centre for Research Excellence in Pulmonary Fibrosis (APP 1099575). We would like to thank the co-ordinators of the Australian IPF registry (AIPFR) in each state for the collection of blood samples. Lung Foundation Australia has established the Australian IPF Registry with the generous support of unrestricted educational grant from Foundation Partners Roche Products Pty. Limited and Boehringer Ingelheim. The writers thank all participants and physicians who contribute to the Registry together with the Manager Sacha Macansh, Coordinators Amy Cashmore, Jessica Bucciarelli, Alysha Riley Karen Symons and Data Manager Faye Janice Lim. We would like to extend our gratitude to Harry Perkins Institute for Medical Research for the use of the ddPCR machine (BioRad). Lastly, we would like to acknowledge the contribution of Alfred Health and the Alfred Health Researchers Jade Jaffar and Glen Westall for supplying the study with frozen fibroblast lines from IPF patients.

References

1. Raghu G, Collard HR, Egan JJ, *et al.* An Official ATS/ERS/JRS/ALAT Statement: Idiopathic Pulmonary Fibrosis: Evidence-based Guidelines for Diagnosis and Management. American Journal of Respiratory and Critical Care Medicine. 2011;183(6):788-824.

2. Ribeiro JR, Lovasco LA, Vanderhyden BC, *et al.* Targeting TBP-Associated Factors in Ovarian Cancer. Frontiers in Oncology. 2014;4:45-.

3. Uhlén M, Fagerberg L, Hallström BM, *et al.* Tissue-based map of the human proteome. Science. 2015;347(6220):1260419.

4. Cividini F, Filoni DN, Pesi R, *et al.* IMP–GMP specific cytosolic 5′-nucleotidase regulates nucleotide pool and prodrug metabolism. Biochimica et Biophysica Acta (BBA)-General Subjects. 2015;1850(7):1354-61.

5. Nance T, Smith KS, Anaya V, *et al.* Transcriptome analysis reveals differential splicing events in IPF lung tissue. PloS One. 2014;9(3):e92111-e.

6. Filoni DN, Pesi R, Careddu MG, *et al.* Initial studies to define the physiologic role of cN-II. Nucleosides, Nucleotides and Nucleic Acids. 2011;30(12):1155-60.

7. Prêle CM, Yao E, O'Donoghue RJ, *et al*. STAT3: a central mediator of pulmonary fibrosis? Proceedings of the American Thoracic Society. 2012;9(3):177-82.

8. Raman M, Earnest S, Zhang K, *et al*. TAO kinases mediate activation of p38 in response to DNA damage. The EMBO journal. 2007;26(8):2005-14.

9. Wu M-F, Wang S-G. Human TAO kinase 1 induces apoptosis in SH-SY5Y cells. Cell Biology International. 2008;32(1):151-6.

10. Yin R, Guo D, Zhang S, *et al.* miR-706 inhibits the oxidative stress-induced activation of PKCα/TAOK1 in liver fibrogenesis. Scientific reports. 2016;6:37509-.

11. Tang Z, Zhang W, Wan C, *et al.* TRAM1 protect HepG2 cells from palmitate induced insulin resistance through ER stress-JNK pathway. Biochemical and Biophysical Research Communications. 2015;457(4):578-84.

12. Burman A, Tanjore H, Blackwell TS. Endoplasmic reticulum stress in pulmonary fibrosis. Matrix Biol. 2018;68-69:355-65.

13. Milligan MJ, Lipovich L. Pseudogene-derived lncRNAs: emerging regulators of gene expression. Frontiers in Genetics. 2015;5(476).

14. Chang W, Wei K, Ho L, *et al.* A critical role for the mTORC2 pathway in lung fibrosis. PloS One. 2014;9(8):e106155.

15. Herazo-Maya JD, Sun J, Molyneaux PL, *et al.* Validation of a 52-gene risk profile for outcome prediction in patients with idiopathic pulmonary fibrosis: an international, multicentre, cohort study. The Lancet Respiratory Medicine. 2017;5(11):857-68.

Competing Interests

Nil.

Funding

This work was funded by a unit grant from the National Health Medical Research Council (APP1066128, APP114776) and Centre for Research Excellence in Pulmonary Fibrosis (CRE-PF), Australia (APP 1099575; 2017-2021).

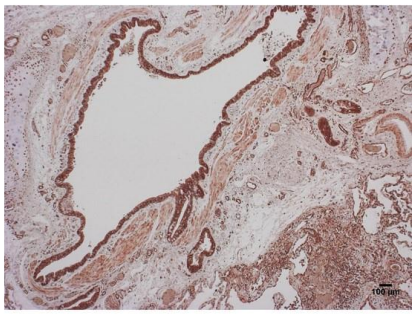
Figure Legends

Figure 1. Differentially expressed genes that distinguish IPF from healthy controls. (A) Comparative concentrations of each gene between IPF vs healthy, progressive vs stable IPF and IPF fibroblasts vs normal fibroblasts. Level of *TAF2* transcript was higher in progressive IPF compared with stable (*p=0.0413). P-values were generated using Kruskal-Wallis test or Mann-Whitney test, adjusted for multiple comparisons with Dunn's multiple comparisons test. Representative examples of 4x TAF2 protein staining in formalin-fixed paraffin embedded IPF lung tissue section (B) and healthy normal lung control tissue (C). Stronger TAF2 expression was observed in the cytoplasm of bronchial epithelial cells, alveolar epithelial cells, smooth muscle cells and fibroblasts in IPF tissue relative to healthy lung.

A.

Gene	Count	IPF (n=57)	Healthy (n=15)	p-value	Progressive (n=24)	Stable (n=33)	p-value
NT5C2	Min	0	0	****	0	4.0	*ns*
	Median	18.9	6.0	<0.0001	24.6	15.0	0.0889
	Max	137.3	13		122.5	137.3	
JAK1	Min	0	0	****	0	3.9	*ns*
	Median	36.0	2.0	<0.0001	37.6	36.0	0.7698
	Max	149.2	10		131.0	149.2	
TAOK1	Min	0	0	****	0	0	*ns*
	Median	17.3	3.0	<0.0001	17.2	17.3	0.6163
	Max	136.9	11.0		136.9	76.9	
TRAM1	Min	0	0	****	0	0	*ns*
	Median	57.9	4.0	<0.0001	61.2	55.7	0.4750
	Max	726.9	14.0		726.9	315.0	
RP11-726G23.6	Min	0	0	****	0	0	*ns*
	Median	10.7	0	<0.0001	12.8	9.3	0.2974
	Max	80.2	28.0		80.2	23.5	
TAF2	Min	0	0	*	0	0	*
	Median	7.5	0	0.0118	12.0	5.5	0.0413
	Max	106.2	27.0		106.0	35.5	
MIR6841	Min	0	0	*	0	0	*ns*
	Median	12.2	7.7	0.0354	10.3	12.3	0.5908
	Max	77.3	36.5		77.3	63.6	

B.



C.