Breath Biopsy® Discovery of Biomarkers for IPF

Luisa Quesada-Arias*[1], Jose Manuel Martinez Manzano*[1], Yichen Chen*[2], Jason Cooper[2], Jason Kinchen[2], Planchart Ferretto[1], Maria Angelica[1], Cheryl Nickerson-Nutter[3], Marc Van der Schee[$2], Ivan Rosas[$1]

*Equal contribution to lead authorship; [$]Equal contribution to senior authorship; [1] Brigham And Women's Hospital, 75 Francis Street, Boston, MA 02115, 617-732-5500, USA; [2] Owlstone Medical, 183 Cambridge Science Park, Milton Road, Cambridge CB4 0GJ, UK; [3] Three Lakes Foundation, Northbrook, Illinois, USA

Abbreviations

- IPF: Idiopathic Pulmonary Fibrosis
- VOCs: volatile organic compounds
- MFs: molecular features
- FEV1: forced expiratory volume in 1 second
- FVC: forced vital capacity
- TLC: total lung capacity
- DLCO: diffusion capacity of the lung for carbon monoxide
- 6MWT: six-minute walking test
- PFT: pulmonary function test

1. Background and objectives

Idiopathic Pulmonary Fibrosis (IPF) is a chronic lung disease of unknown cause associated with the development of progressive and irreversible fibrosis of the lung parenchyma. IPF affects approximately 3 million people worldwide[1], with a mean survival rate of 2-3 years after diagnosis[2].

Hypothesis:

Metabolic changes associated with IPF produce unique volatile organic compounds (VOCs) detectable on breath that could be applied as diagnostic biomarkers.

Study Outcomes:

1. Validate recently-described findings previously associated with IPF[3]
2. Discover candidate molecular features (MFs) that can discriminate subjects with IPF from healthy control
3. Evaluate association of these MFs with pulmonary function tests (PFTs)

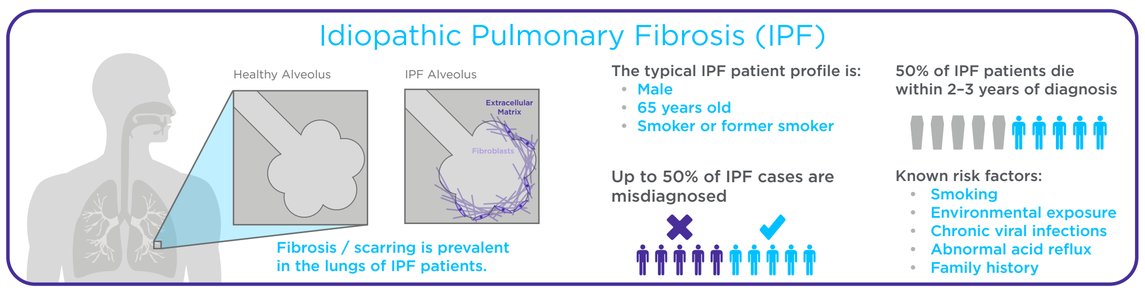

Figure 1: A selection of information and statistics relating to IPF[4].

2. Methods

Subjects were enrolled at Brigham and Women's Hospital, Boston, MA; study demographics are shown in **Table 1**.

Breath Biopsy samples were collected using the ReCIVA® Breath Sampler (**Figure 2**), developed by Owlstone Medical, and analyzed with thermal desorption gas chromatography mass spectrometry (TD-GC-MS) using the Breath Biopsy Platform.

Type	Variable	Control (n=47)	IPF (n=58)	p-value
Demographic	Age (years)	60.0 ± 6.00	72.0 ± 4.00	$1.04x10^{-10}$
	Gender, male	17 (36.2%)	33 (56.9%)	0.0491
	Race, white	47 (100%)	54 (93.1%)	0.126
	Ever Smoked	11 (23.4%)	36 (62.1%)	0.000812
	BMI	27.7 ± 3.60	26.8 ± 2.98	0.0629
PFT	FEV1 - % of predicted value	107 ± 11.0	84.0 ± 12.5	$1.36x10^{-9}$
	FVC - % of predicted value	108 ± 9.00	79.0 ± 13.5	$3x10^{-13}$
	TLC - % of predicted value*	104 ± 9.00	73.0 ± 11.0	$3.3x10^{-7}$
	DLCO - % of predicted value**	89.0 ± 12.0	48.0 ± 10.0	$2.98x10^{-10}$
	6MWT - meters walked***	483 ± 69.0	414 ± 46.0	$2.85x10^{-5}$

Table 1: *Baseline Characteristics of IPF and Control participants. P-values are derived from Mann-Whitney U-test (continuous variables) or Fisher's Exact test (categorical variables). No multiple testing correction was applied. Median is shown ± median absolute deviation (MAD). Categorical variables reported in frequency and percentage under parenthesis. P-value < 0.001 is highlighted in bold. Data missing in the following categories, all of them had pre-6MWT breath samples: *TLC in 9 IPF subject **DLCO in 6 IPF subject. *** Meters Walked in 2 IPF subject.*

Figure 2: *Overview of breath sample acquisition and analysis. Left section pictures the Breath Biopsy® Collection Station, consisting of ReCIVA® Breath Sampler, CASPER™ Portable Air Supply and Breath Biopsy Collect Software.*

4. Conclusions and Further Work

Here, we have confirmed the association between MF38 (acetoin) and IPF status. The discovery analysis identified 22 molecular features (MFs) with unadjusted p < 0.05 when comparing IPF and healthy control using breath collected via Owlstone's BB technology. A biomarker selection paradigm (multivariate analysis) identified 4 candidates that provide non-redundant information on IPF status. Differential patterns of associations suggest these candidates may report on orthogonal aspects of the disease process.

As a path forward, results presented in this pilot study should be replicated in a follow-on study with similar design. Extension of these findings, particularly intent-to-diagnose cohorts (e.g. subjects with idiopathic lung abnormalities), would provide a valuable test around the clinical utility of biomarker findings.

3. Results

Following feature extraction and quality control, a total of 129 molecular features (MFs) from 124 subjects were retained in the discovery dataset. The analysis was performed with the exhaled breath samples from 58 (45.6%) subjects with IPF and 47 (37%) control subjects.

We also evaluated four VOCs previously associated with breath in the literature. Chemical standards were used to target identification of as p-cymene, acetoin, isoprene, and ethylbenzene. MF38 (acetoin) showed an association when IPF was compared to control (P-value = 0.00699, one-sided test) using univariate logistic regression (**Figure 3**); remaining candidates did not show sufficient evidence of association and were excluded from subsequent analysis. Previous work has speculated that acetoin may be associated with exposure to ROS, but could also derived from environmental exposure (ie, diet).

A

Molecular Feature	Coefficient, MF term	Raw one-tailed p-value
MF38 (acetoin)	0.756 ± 0.307	**0.00699**
MF54 (Ethylbenzene)	-0.401 ± 0.266	0.933
MF95 (p-cymene)	-0.349 ± 0.282	0.108
MF4 (isoprene)	0.226 ± 0.263	0.195

B

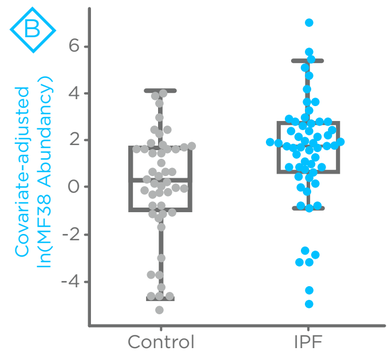

Figure 3: *A. Results from IPF-Control univariate analysis for validating VOCs found in a previous study[3]. Coefficient is shown ± standard error. Tentative ID is shown in parentheses for each MF; p < 0.05 shown in bold B. MF38 (acetoin) from previous findings[3] was validated in IPF-Control analysis (uncorrected P-value = 0.00699, one-side test).*

In the univariate discovery analysis, twenty-two MFs had uncorrected P-value < 0.05 in pre-6MWT univariate IPF-Control analysis (**Table 2**). An evaluation of tentative IDs for these compounds (not shown) suggested a diverse family of compounds that may reflect an impact of structural changes in the lung on VOC diffusion from the blood.

Molecular Feature	Coefficient, MF term	Raw two-tailed p-value	Multiple-testing corrected P-value
MF121	1.090 ± 0.351	0.00201	0.074
MF33	1.070 ± 0.349	0.00207	0.074
MF46	1.070 ± 0.362	0.00313	0.074
MF96	-0.911 ± 0.314	0.00373	0.074
MF75	1.050 ± 0.369	0.00439	0.074
MF73	-1.220 ± 0.431	0.00453	0.074

Table 2: *Subset of MFs with uncorrected P-value < 0.05 in IPF-Control univariate analysis. Benjamini-Hochberg method was applied for multiple testing correction.*

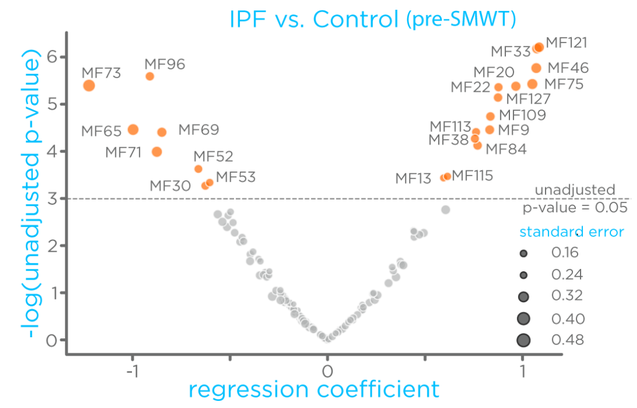

Figure 3: *Volcano plot of molecular features from pre-SMWT samples plotted by -log (unadjusted p-value) and regression coefficient between IPF cases and controls. Features are colored by significance thresholds and sized by standard error.*

In multivariable discovery analysis, logistic regression with L1 regularization was conducted on data corrected for multicollinearity to identify biomarker candidates that jointly are informative in predicting IPF status (resulting candidates shown in **Table 3**). These features and MF38 (acetoin) were then tested for association with physiological parameters (FEV1, FCV, DLCO and TLC, see **Table 3**).; MF38 showed significant negative association with FEV1 (P-value=0.00919). The three informative MF predictors identified in IPF-Control showed consistent direction of association with some physiological parameters in the univariate analyses: MF46 was found to have significant negative association with FEV1, FVC and TLC (P-values = 0.00485, 0.00409 and 0.0125, respectively). MF96 with significant positive association with DLCO (P-value = 0.0153). And lastly, MF121 showed significant negative association with FEV1, FVC, DLCO and TLC (P-value = 0.0037, 0.000764 , 0.00415 and 0.0208, respectively).

Molecular Feature	FEV1		FCV		DLCO		TLC	
	Coefficient, MF term	Raw two-tailed p-value	Coefficient, MF term	Raw two-tailed p-value	Coefficient, MF term	Raw two-tailed p-value	Coefficient, MF term	Raw two-tailed p-value
MF38	-4.88 ± 1.84	0.00919	-3.31 ± 1.94	0.0891	-2.38 ± 1.80	0.019	-1.25 ± 1.80	0.487
MF46	-5.26 ± 1.83	0.00485	-5.62 ± 1.92	0.00409	-2.17 ± 1.84	0.239	-4.50 ± 1.77	0.0125
MF96	3.17 ± 1.87	0.0934	3.11 ± 1.99	0.12	5.83 ± 1.79	0.00153	2.81 ± 1.79	0.119
MF121	-5.43 ± 1.83	0.0037	-6.4 ± 1.85	0.000764	-5.09 ± 1.74	0.00415	-4.12 ± 1.75	0.0208

Table 3: *FVC, FEV, DLCO and TLC - univariate results for MF38, MF46, MF96 and MF121. Coefficient is shown ± standard error.*

5. References

1. Nalysnyk L., et al., Incidence and Prevalence of Idiopathic Pulmonary Fibrosis: Review of the Literature. Eur, *Respir Rev.* 2012; 21 (126): 355-361.
2. Raghu G., et al., An Official ATS/ERS/JRS/ALAT Statement: Idiopathic Pulmonary Fibrosis: Evidence-based, Guidelines for Diagnosis and Management. *Am J Respir Crit Care Med.* 2011; 183: 788–824.
3. Yamada YI, Yamada G, Otsuka M, et al. Volatile Organic Compounds in Exhaled Breath of Idiopathic Pulmonary Fibrosis for Discrimination from Healthy Subjects. *Lung*. Published online 2017. doi:10.1007/s00408-017-9979-3
4. Boehringer Ingelheim, https://www.boehringer-ingelheim.com/respiratory/ipf/ipf-video-infographic

Further Resources

Breath Biopsy: The Complete Guide (3rd Edition)
owlstonemedical.com/breath-biopsy-guide

Breath Biopsy for Respiratory Disease webinar
owlstonemedical.com/respiratory-webinar

Breath Biopsy Products & Services
owlstonemedical.com/products