Interaction between the promoter *MUC5B* polymorphism and mucin expression: is there a difference according to ILD subtype?

Carmel J. Stock[1,], Caterina Conti[1,2], Angeles Montero-Fernandez[3], Gaetano Caramori[4], Philip L. Molyneaux[1], Peter M. George[1], Maria Kokosi[1], Vasilis Kouranos[1], Toby M. Maher[1], Felix Chua[1], Alexandra Rice[3], Christopher P. Denton[5], Andrew G. Nicholson[3], Athol U. Wells[1], Piersante Sestini[6], Elisabetta Renzoni[1]

[1]Interstitial Lung Disease Unit, Royal Brompton and Harefield NHS Foundation Trust/ National Heart and Lung Institute, Imperial College, London, UK

[2]Respiratory Unit ASST-Ospedale Papa Giovanni XXIII, Bergamo, Italy

[3]Department of Histopathology, Royal Brompton and Harefield NHS Foundation Trust/ National Heart and Lung Institute, Imperial College, London, UK

[4]Pneumologia, Dipartimento di Scienze Biomediche, Odontoiatriche e delle Immagini Morfologiche e Funzionali (BIOMORF), Università degli Studi di Messina, Messina, Italy

[5]Centre for Rheumatology and Connective Tissue Diseases, Royal Free and University College Medical School, London, UK

[6] Department of Medical and Surgical Sciences & Neurosciences, University of Siena, Siena, Italy

Corresponding author:
Dr Carmel Stock, Interstitial Lung Disease Unit, National Heart and Lung Institute, Imperial College London, Royal Brompton and Harefield NHS Foundation Trust, Sydney Street, London, UK, SW3 6NP email: c.stock@imperial.ac.uk Telephone: 0207 351 8456

Take home message (256 characters):
MUC5B rs35705950 T-carrier status is associated with increased protein expression of MUC5B in distal airways only in IPF, but not in controls, I-NSIP or SSc-NSIP. T-carrier status appears to be associated with a better prognosis and less severe disease.

Word count: 992

Abstract: (99 words)

The *MUC5B* promoter variant rs35705950 is associated with idiopathic pulmonary fibrosis (IPF). MUC5B glycoprotein is overexpressed in IPF lungs. We examined immunohistochemical expression of MUC5B in different ILD patterns according to rs35705950 T-allele carriage. We observed increased expression of MUC5B in T-allele carriers in both distal airways and honeycomb cysts in patients with IPF (N=23), but no

difference in MUC5B expression according to T-carrier status in the distal airways of patients with idiopathic non-specific interstitial pneumonitis (I-NSIP) (N=18), in scleroderma-associated NSIP (N=15) or in control lungs (N=20), suggesting that tissue overexpression in *MUC5B* rs35705950 T-carriers is specific to IPF.

The *MUC5B* promoter variant rs35705950 is the strongest genetic risk factor for idiopathic pulmonary fibrosis (IPF), and may also be associated with a less rapid disease course.[1,2] How *MUC5B* overexpression contributes to development of fibrosis remains poorly understood. We previously described overexpression of MUC5B in the small airways and honeycomb cysts of patients with IPF compared to controls. We sought to establish whether the overexpression seen in IPF was linked to the *MUC5B* allele, and whether differences in the relationship between allele and lung MUC5B expression were observed among different ILD patterns.

The current study utilises MUC5B immunohistochemical expression data, analysed as previously described.[3] Briefly, quantification of MUC5B expression was evaluated in available sequential histology blocks from 2003-2010 of surgical lung biopsies from 23 patients with IPF/usual interstitial pneumonia (UIP), 18 with idiopathic non-specific interstitial pneumonia (I-NSIP), 15 with scleroderma (SSc)-NSIP, and 20 normal lung tissue peripheral to resected cancer. In each biopsy, three distal airways, and in UIP, three honeycomb cysts, were evaluated. In each area, quantification of the proportion of MUC5B+ cells was evaluated in six randomly selected fields, each containing 100 airway epithelial cells (or honeycomb cyst).[3] DNA was extracted from formalin-fixed paraffin-embedded (FFPE) tissue using the QIAamp DNA FFPE tissue kit (Qiagen). Genotyping was performed using a commercially available TaqMan® assay (Applied Biosystems).

Comparison of the genotype distribution between the groups was performed using logistic regression and expressed as Odds Ratio (OR) ±Standard Error (SE). To compare MUC5B immunohistochemical expression between groups, multilevel mixed-effects linear regression was performed, in which patients were analysed as random effect variables with airways (or honeycomb cysts) and microscopic fields nested into patients, and with the binary variable *MUC5B* rs35705950 T allele carriage (hereafter T-carrier) and diagnostic group, and their interaction, as fixed effect variables. Differences reaching statistical significance in the complete models were then tested individually using simplified models omitting irrelevant variables. Impact of the T-carrier and the percentage of MUC5B+ cells on the composite physiologic index (CPI= 91.0 - (0.65xDLCO%predicted) - (0.53xFVC%) + (0.34xFEV1%) as a measure of disease severity[4] was analysed using generalised linear models with the CPI as the dependent variable and the T-carrier and/or mean percentage of MUC5B+ cells per patient as independent variables. A p value ≤0.05 was considered significant (Stata 15).

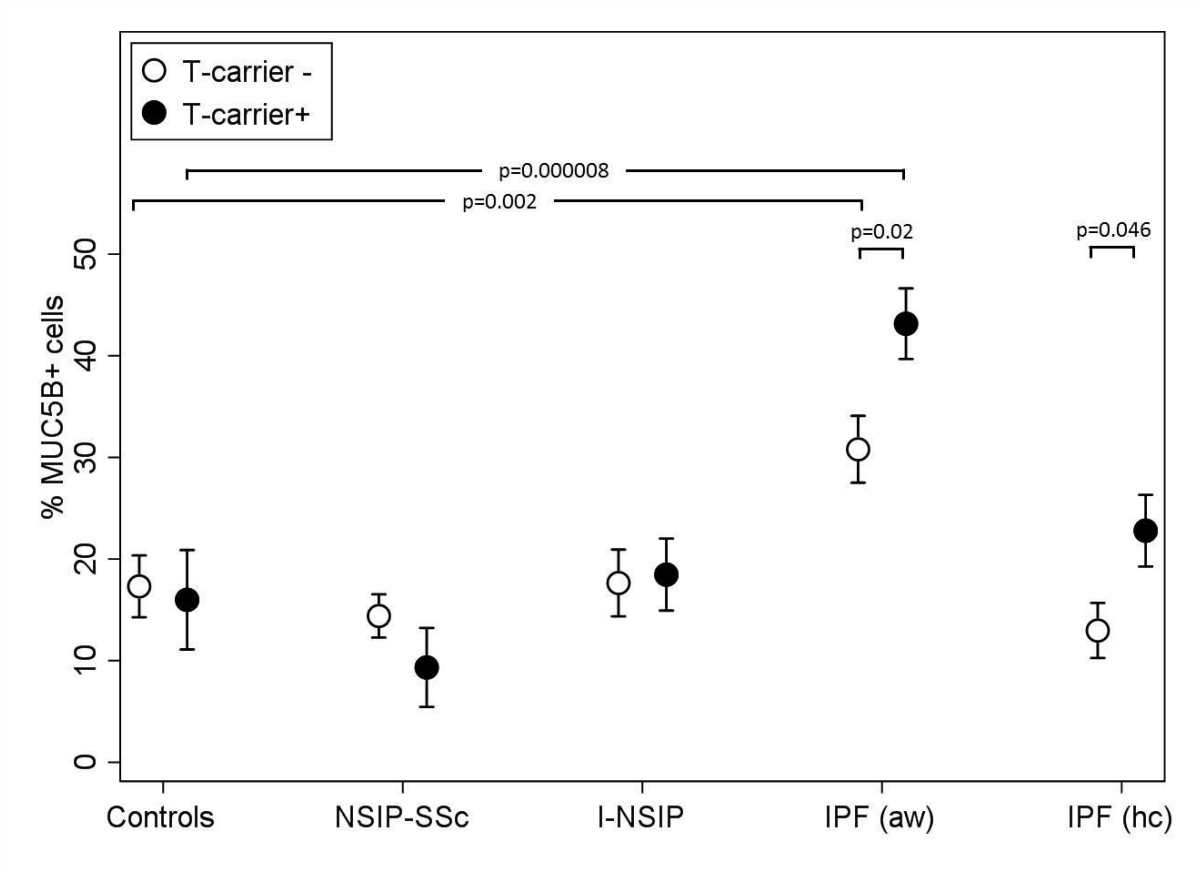
Analysis of the *MUC5B* genotype and allele carriage revealed, as expected, a significantly higher proportion of T-carriers in IPF patients compared to controls (OR: 5.3 ±3.6, p=0.015) but no difference between the other groups and controls (data not shown). To assess the

effect of T-carrier on MUC5B tissue expression in the different diagnostic groups, the differences in the mean percentage of MUC5B+ cells in the distal airways of the four groups (and for IPF also the honeycomb cysts) according to T-carrier status was assessed**.** As shown in Figure 1, a significantly increased expression of MUC5B in the distal airways compared to controls was observed only in IPF patients, regardless of T-carrier positivity. Furthermore, an increase in MUC5B expression in distal airways was associated with T-carrier positivity only in IPF (p=0.02 vs T-carrier negative), resulting in a greater difference with controls in T-carrier positive than in T-carrier negative patients (p=0.000008 between T-carrier positive, and p=0.002 between T-carrier negative patients and controls). An increased expression according to T-carrier positivity in IPF patients was also seen in honeycomb cysts (p=0.047). There was instead no significant difference in MUC5B expression according to T-carrier status, in I-NSIP, SSc-NSIP, or in controls.

We then focused on assessing the relationship of MUC5B and T-carrier status with disease severity in IPF patients. The CPI score was higher (more severe disease) in T-carrier negative IPF patients (mean±SE: 48.1±2.8), compared to T-carrier positive IPF patients (38.2±2.7, p<0.001). The CPI score was inversely correlated to the percentage of MUC5B+ cells (slope: -0.28±0.11, p=0.011). When both variables were included in the multivariable model, only the T-carrier effect remained significant (slope: -8.2±2.9,p=0.005), while the effect of the MUC5B+ cells was no longer significant (slope: -0.14±0.08, p=0.08).

We conclude that *MUC5B* rs35705950 T-carrier status is associated with increased expression of MUC5B, but only in IPF. We also find that the increase in MUC5B+ cells in IPF is not solely related to the higher frequency of T-carriers, but also occurs independently of allele carriage, as described for mRNA expression by Seibold et al.[1] Interestingly, Seibold et al. reported significantly higher mRNA expression in T-carriers in both control and IPF lungs. Nakano et al. reported a strong correlation between MUC5B promoter activity and MUC5B+ epithelial cells in IPF lungs, but had not investigated other ILD patterns.[5] An explanation for the specificity of the increased MUC5B staining in T-carriers observed only for IPF lungs in this study will require further investigation. Downstream 32bp of rs35705950 is a highly conserved FOXA2 binding motif. This region is hypermethylated in the presence of IPF, increased expression of MUC5B in lung tissue, and the rs35705950 risk allele. This hypermethylation may result in increased occupancy of FOXA2 in the binding motif, leading to increased MUC5B expression.[6] Although analysis of mRNA expression would have been desirable, mRNA in FFPE tissue can be heavily degraded, resulting in severe limitations in the reliability of the relative mRNA expression in these samples,[7,8,9] and fresh frozen samples, known to provide a better preservation of mRNA, were not available. The *MUC5B* rs35705950 T allele also seems to be associated with less severe disease. Although this may be related to a separate mechanism than increased MUC5B expression, further larger studies are needed to assess relative contributions of the rs35705950 T-allele and MUC5B glycoprotein expression in relation to IPF severity, and explore the mechanisms underlying this association. Our study did not have sufficient power to reliably assess links between the T-allele and MUC5B tissue expression and prognosis. In conclusion, we find that a positive relationship between the MUC5B risk allele and expression of MUC5B glycoprotein in the lungs is specific to IPF. Further studies are needed to confirm this observation and explore its underlying mechanisms.

Figure 1. Comparison of distal airway MUC5B expression between control lungs, NSIP-associated with SSc (NSIP-SSc), idiopathic NSIP (I-NSIP) and IPF. In IPF both distal airways (aw) and honeycomb cysts (hc) were evaluated. MUC5B expression was significantly higher in IPF distal airways, but not in honeycomb cysts compared to controls, while no difference was observed between either I-NSIP or SSc-NSIP and controls. T allele carriage was associated with signifcantly higher MUC5B expression both in IPF airways and honeycomb cyst, while no difference was observed in the other groups (Scheffe's post hoc analysis).



Reference List

1. Seibold M.A., Wise A.L., Speer M.C. et al. A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503-1512.

2. Peljto A.L., Zhang Y., Fingerlin T.E. et al. Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013;309:2232-2239.

3. Conti C., Montero-Fernandez A., Borg E. et al. Mucins MUC5B and MUC5AC in Distal Airways and Honeycomb Spaces: Comparison among Idiopathic Pulmonary Fibrosis/Usual Interstitial Pneumonia, Fibrotic Nonspecific Interstitial Pneumonitis, and Control Lungs. *Am J Respir Crit Care Med* 2016;193:462-464.

4. Wells A.U., Desai S.R., Rubens M.B. et al. Idiopathic pulmonary fibrosis: a composite physiologic index derived from disease extent observed by computed tomography. *Am J Respir Crit Care Med* 2003;167:962-969.

5. Nakano Y., Yang I.V., Walts A.D. et al. MUC5B Promoter Variant rs35705950 Affects MUC5B Expression in the Distal Airways in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2016;193:464-466.

6. Helling B.A., Gerber A.N., Kadiyala V. et al. Regulation of MUC5B Expression in Idiopathic Pulmonary Fibrosis. *Am J Respir Cell Mol Biol* 2017;57:91-99.

7. Macabeo-Ong M., Ginzinger D.G., Dekker N. et al. Effect of duration of fixation on quantitative reverse transcription polymerase chain reaction analyses. *Mod Pathol* 2002;15:979-987.

8. Srinivasan M., Sedmak D., and Jewell S. Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *Am J Pathol* 2002;161:1961-1971.

9. von A.S., Missel A., Bendrat K. et al. Determinants of RNA quality from FFPE samples. *PLoS One* 2007;2:e1261-