Telomere length and genetic variant associations with interstitial lung disease progression and survival

Chad A. Newton[1], Justin M. Oldham[2], Brett Ley[3], Vikram Anand[1], Ayodeji Adegunsoye[4], Gabrielle Liu[3], Kiran Batra[5], Jose Torrealba[6], Julia Kozlitina[7], Craig Glazer[1], Mary E. Strek[4], Paul J. Wolters[3], Imre Noth[8] and Christine Kim Garcia[1,7]

Affiliations: [1]Dept of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, TX, USA. [2]Dept of Internal Medicine, University of California at Davis, Davis, CA, USA. [3]Dept of Medicine, University of California San Francisco, San Francisco, CA, USA. [4]Dept of Medicine, University of Chicago, Chicago, IL, USA. [5]Dept of Radiology, University of Texas Southwestern Medical Center, Dallas, TX, USA. [6]Dept of Pathology, University of Texas Southwestern Medical Center, Dallas, TX, USA. [7]Eugene McDermott Center for Human Growth and Development, University of Texas Southwestern Medical Center, Dallas, TX, USA. [8]Dept of Medicine, University of Virginia School of Medicine, Charlottesville, VA, USA.

Correspondence: Chad A. Newton, 5323 Harry Hines Blvd, Dallas, TX 75390-8558, USA.
E-mail: chad.newton@utsouthwestern.edu

@ERSpublications
Leukocyte telomere length and *MUC5B* minor allele frequency are similar for IPAF and the combined CTD-ILD group; however, the associations between these genomic markers and clinical outcomes are different for these two types of ILD http://ow.ly/wXem30njRkg

Cite this article as: Newton CA, Oldham JM, Ley B, *et al*. Telomere length and genetic variant associations with interstitial lung disease progression and survival. *Eur Respir J* 2019; 53: 1801641 [https://doi.org/10.1183/13993003.01641-2018].

ABSTRACT Leukocyte telomere length (LTL), *MUC5B* rs35705950 and *TOLLIP* rs5743890 have been associated with idiopathic pulmonary fibrosis (IPF).

In this observational cohort study, we assessed the associations between these genomic markers and outcomes of survival and rate of disease progression in patients with interstitial pneumonia with autoimmune features (IPAF, n=250) and connective tissue disease-associated interstitial lung disease (CTD-ILD, n=248). IPF (n=499) was used as a comparator.

The LTL of IPAF and CTD-ILD patients (mean age-adjusted log-transformed T/S of −0.05±0.29 and −0.04±0.25, respectively) is longer than that of IPF patients (−0.17±0.32). For IPAF patients, LTL <10th percentile is associated with faster lung function decline compared to LTL ⩾10th percentile (−6.43% per year *versus* −0.86% per year; p<0.0001) and worse transplant-free survival (hazard ratio 2.97, 95% CI 1.70–5.20; p=0.00014). The *MUC5B* rs35705950 minor allele frequency (MAF) is greater for IPAF patients (23.2, 95% CI 18.8–28.2; p<0.0001) than controls and is associated with worse transplant-free IPAF survival (hazard ratio 1.92, 95% CI 1.18–3.13; p=0.0091). Rheumatoid arthritis (RA)-associated ILD (RA-ILD) has a shorter LTL than non-RA CTD-ILD (−0.14±0.27 *versus* −0.01±0.23; p=0.00055) and higher *MUC5B* MAF (34.6, 95% CI 24.4–46.3 *versus* 14.1, 95% CI 9.8–20.0; p=0.00025). Neither LTL nor *MUC5B* are associated with transplant-free CTD-ILD survival.

LTL and *MUC5B* MAF have different associations with lung function progression and survival for IPAF and CTD-ILD.

This article has supplementary material available from erj.ersjournals.com.

Received: Aug 28 2018 | Accepted after revision: Dec 27 2018

Introduction

The interstitial lung diseases (ILDs) are a heterogeneous group of disorders characterised by fibrosis of the lung. Determining the discrete ILD diagnosis for each patient based on clinical, radiographic and histopathologic features is critically important for informing prognosis. Idiopathic pulmonary fibrosis (IPF) is the prototypical fibrosing lung disease that has a progressive and lethal course with median survival of approximately 3 years [1], as opposed to connective tissue disease-associated interstitial lung disease (CTD-ILD) which has a much more favourable prognosis. Specific ILD diagnoses also inform treatment decisions; for example, exposure to immunosuppressive medications is associated with worse outcomes in IPF [2] but may be beneficial for CTD-ILD [3–5]. Some patients exhibit clinical features that overlap those of IPF and CTD-ILD. Recently, a joint European Respiratory Society (ERS) and American Thoracic Society (ATS) task force proposed criteria to facilitate recognition and study of this ILD subtype, termed "interstitial pneumonia with autoimmune features (IPAF)". The criteria outline clinical, serologic and morphologic features suggestive of an underlying autoimmune disease in the absence of extra-pulmonary manifestations of a well-defined connective tissue disease (CTD) [6]. While studies have described the clinical features and survival characteristics of patients with IPAF [7], little is known about the genetic determinants of clinical outcomes in this population.

Genetic and genomic factors are associated with the risk of developing ILD and influence clinical outcomes. Common variants such as single nucleotide polymorphisms (SNPs) in the *MUC5B* and *TOLLIP* genes are more common in IPF patients compared to controls [8–10]. These SNPs can inform mortality risk and rate of disease progression [10–12]. Pathogenic rare variants in telomere-maintenance genes have been linked to pulmonary fibrosis and shortened telomeres, the protective ends of chromosomes. Patients with telomere-related rare variants in *TERT*, *TERC*, *PARN* or *RTEL1* can manifest many forms of pulmonary fibrosis including IPF, IPAF and CTD-ILD, but uniformly exhibit relentless disease progression and poor survival [13]. Shortened age-adjusted leukocyte telomere length (LTL) has also been associated with worse survival in patients with IPF [14–16] and chronic hypersensitivity pneumonitis (CHP) [17].

The objective of this study was to determine if specific genetic and genomic markers associated with survival in IPF are also associated with survival and rate of disease progression in patients with IPAF and CTD-ILD. Genotypes of the *MUC5B* rs35705950 and *TOLLIP* rs5743890 SNPs, as well as peripheral blood LTLs were measured across independent cohorts of patients. IPF patients were included as a comparator group.

Methods

Study design and populations

This retrospective cohort study included all patients with a diagnosis of IPAF, CTD-ILD, or IPF who were enroled in longitudinal registries at three academic medical centres. Patients were enroled at the University of Texas Southwestern (Dallas, TX, USA; the UTSW cohort; June 17, 2003 to July 01, 2017), the University of California San Francisco (San Francisco, CA, USA; the UCSF cohort; November 14, 1998 to September 25, 2017) and the University of Chicago (Chicago, IL, USA; the Chicago cohort; January 24, 2006 to September 01, 2017). Each participant provided written informed consent and a peripheral blood sample at enrolment into the respective registries. Multidisciplinary discussion informed diagnosis at each site independently. IPF was diagnosed according to consensus guidelines [1] and the CTD-ILD diagnoses included rheumatologic evaluation. Each site retrospectively identified patients who met classification criteria for IPAF. The IPAF classification required at least one criterion from two or more domains (clinical, serologic, or morphologic) [6]. In order to maintain consistency of the IPAF diagnosis across sites, unexplained intrinsic airway disease was not considered a component of the morphologic criteria for current or prior smokers. In addition, pulmonary vasculopathy required mean pulmonary arterial pressure (PAP) >25 mmHg and wedge pressure <15 mmHg on right-heart catheterisation, or estimated right-ventricular systolic pressure >40 mmHg by echocardiography, or presence of vasculopathy on a histopathologic specimen. A thoracic radiologist and a thoracic pathologist at each site reviewed high-resolution computed tomography (HRCT) scans of the chest and available pathologic specimens to confirm the presence of IPAF features. Clinical information including demographics, symptoms, signs, laboratory results and longitudinal pulmonary function tests (PFTs) were abstracted from medical records (ethnicity was self-reported). This study was approved by the institutional review boards of the University of Texas Southwestern Medical Center, the University of California San Francisco and the University of Chicago for their respective cohorts. The majority of IPF patients (UTSW (n=149), UCSF (n=54) and Chicago (n=139)) [14] and the IPAF patients from Chicago (n=112) were included in separate previous studies [7].

Genotyping and telomere length measurements

Genomic DNA was isolated from peripheral blood leukocytes using an Autopure LS instrument (the UTSW cohort), a Gentra Puregene Blood kit (the UCSF cohort), or a Flexigene DNA kit (the Chicago cohort) (all from Qiagen, Valencia, CA, USA). LTL was measured for the UTSW cohort using a

quantitative polymerase chain assay [14, 18, 19] and for the UCSF and Chicago cohorts using an identical protocol except that each sample was diluted to 20 $ng \cdot \mu L^{-1}$ instead of 50 $ng \cdot \mu L^{-1}$ before its addition to the PCR reaction. Age-adjusted LTL was calculated using normal controls and presented as observed minus expected values. The intraclass correlation for the LTL measurement was 0.987 (95% CI 0.983–0.991), 0.989 (95% CI 0.982–0.994) and 0.940 (95% CI 0.924–0.953) for the UTSW, UCSF and Chicago cohorts, respectively.

SNP genotyping was performed with the UTSW cohort for *MUC5B* rs35705950 and *TOLLIP* rs5743890 using the Taqman SNP Genotyping Assay (Applied Biosystems, Foster City, CA, USA). The SNP genotype minor allele frequency (MAF) was calculated along with binomial 95% CIs and reported for patients of non-Hispanic white ethnicity and compared to controls from the European population of the 1000 Genomes Project Phase 3 (project 1) [20].

Statistics

Categorical variables were expressed as counts and percentages and were compared across groups using the Chi-squared test when the expected count for each cell was five or more (otherwise Fisher's exact test was used). Continuous variables were expressed as means with standard deviations and were compared using the two-tailed t-test (for two group comparisons) or one-way ANOVA (for more than two group comparisons). For comparisons across more than two groups, *post hoc* analysis was performed using pairwise comparisons with Bonferroni adjustment.

The primary outcome of this study was transplant-free survival for patients with IPAF and CTD-ILD, defined as time from enrolment to death or transplant. Overall survival, with censoring at the time of transplant, was evaluated as the secondary endpoint in sensitivity analysis. The association between genomic predictors and the primary and secondary endpoints was tested using multivariate Cox proportional hazards regression models stratified by cohort. The genomic predictors for the primary and secondary analyses included the *MUC5B* rs35705950 and *TOLLIP* rs5743890 genotypes (homozygous wild-type *versus* heterozygous and homozygous minor alleles), as well as age-adjusted LTL (<10th or ⩾10th percentile), as previous studies have shown this to be an informative cut-off point [14, 17–19]. To account for baseline differences and known confounders, the association between transplant-free survival and each genomic predictor was adjusted for age, gender, ethnicity, baseline % predicted forced vital capacity (FVC) and baseline % predicted diffusing capacity of lung for carbon monoxide (D_{LCO}) without imputation for missing data. An additional model was evaluated that included the pattern of ILD (usual interstitial pneumonia (UIP), yes/no) along with age, gender, ethnicity, baseline % predicted FVC and baseline % predicted D_{LCO}, to determine if the pattern of ILD influenced the genomic marker associations with transplant-free survival. Both LTL and *MUC5B* rs35705950 were included as variables in a multivariable model to assess their independent associations with transplant-free survival. A Bonferroni adjusted alpha of 0.017 (0.05/3) was used as the significance threshold to account for multiple testing with three genomic predictors (LTL, *MUC5B* and *TOLLIP*) per diagnosis for the transplant-free and overall survival analyses. There was no evidence of non-proportional hazards noted by plotting scaled Schoenfeld residuals against time for each covariate included in the Cox models.

To quantify the rate of disease progression, we assessed the change in % predicted FVC per year using linear mixed-effects models including patients with three or more available measurements that spanned ⩾90 days. Age, gender, ethnicity and smoking status were included as fixed effects in the model to account for baseline differences. The changes in % predicted FVC per year were reported for each genomic categorical predictor. The parameters were estimated using the restricted maximum likelihood procedure. The need for random effects was assessed using likelihood ratio tests and random slopes and intercepts were included in the model. A Bonferroni adjusted alpha of 0.017 (0.05/3) was used as the significance threshold to account for multiple testing similar to the survival analysis. All p-values less than 0.05 were considered significant unless otherwise stated. All analyses were performed using R statistical analysis software, version 3.3.2 (The R Project for Statistical Computing; www.r-project.org).

Results

Characteristics of disease groups

This study included 250 patients with IPAF (UTSW cohort (n=73), UCSF cohort (n=63), Chicago cohort (n=114)), 248 patients with CTD-ILD (UTSW cohort (n=102), Chicago cohort (n=146)) and 499 patients with IPF (UTSW cohort (n=303), UCSF cohort (n=54), Chicago cohort (n=142)) (table 1). Differences among the cohorts collected from the independent sites are listed in supplementary tables S1–S3. Overall, the demographic characteristics (age, gender and ethnicity) of the IPAF cohort fell between the IPF and CTD-ILD cohorts. The most common CTD subtypes represented in the combined CTD-ILD cohort were scleroderma (SSc; 74 out of 248 patients (30%)) and rheumatoid arthritis (RA; 62 out of 248 patients (25%)).

TABLE 1 Characteristics of patients with idiopathic pulmonary fibrosis (IPF), interstitial pneumonia with autoimmune features (IPAF) and connective tissue disease-associated interstitial lung disease (CTD-ILD)

Characteristic	IPF (n=499)	IPAF (n=250)	CTD-ILD[#] (n=248)	p-value[¶]
Age years	65.7±9.6	60.5±11.1	53.8±13.4	<0.0001
Male gender	368 (74)	112 (45)	70 (28)	<0.0001
Ethnicity				<0.0001
Non-Hispanic white	437 (87)	170 (68)	138 (56)	
Hispanic or Latino	34 (7)	26 (10)	30 (12)	
Black	17 (4)	36 (14)	75 (30)	
Asian	6 (1)	12 (5)	5 (2)	
Other or unknown	5 (1)	6 (2)	0 (0)	
Ever smoker	317 (66)	134 (54)	105 (42)	<0.0001
Family history	61 (20)	10 (4)	7 (3)	<0.0001
PFT				
FVC % predicted	67±18 (n=418)	64±19 (n=228)	68±19 (n=214)	0.08
*D*LCO % predicted	47±17 (n=386)	48±18 (n=212)	53±20 (n=197)	0.001
Telomere length[+]				
Observed−expected	−0.17±0.32	−0.05±0.29	−0.04±0.25	<0.0001[§]
<10th percentile	156 (31)	40 (16)	32 (13)	<0.0001[§]
SNPs[f]				
MUC5B rs35705950 MAF	34.2 (95% CI 31.1–37.5) (n=437)	23.2 (95% CI 18.8–28.2) (n=166)	19.9 (95% CI 15.5–25.2) (n=138)	<0.0001[##]
TOLLIP rs5743890 MAF	12.4 (95% CI 10.3–14.8) (n=437)	15.0 (95% CI 11.4–19.5) (n=163)	14.2 (95% CI 10.4–19.1) (n=137)	0.42
Follow-up years (median (IQR))	2.97 (1.54–4.86)	2.86 (1.25–3.71)	4.60 (1.88–8.21)	<0.0001
Disease progression [¶¶]				
ΔFVC % predicted·year^{-1}	−5.37 (95% CI −6.10 to −4.66) (n=212)	−1.80 (95% CI −2.70 to −1.0) (n=163)	−0.64 (95% CI −0.99 to −0.30) (n=181)	<0.0001
Survival				
Transplant-free survival years (median)	3.75 (95% CI 3.48–4.40)	5.61 (95% CI 4.88–7.07)	11.88 (95% CI 9.18–NA)	<0.0001

Data are presented as n (%) or mean±SD, unless otherwise stated. PFT: pulmonary function test; FVC: forced vital capacity; *D*LCO: diffusing capacity of the lung for carbon monoxide; SNP: single nucleotide polymorphism; MAF: minor allele frequency; IQR: interquartile range; NA: not available. [#]: CTD-ILD diagnoses include scleroderma (SSc; n=74), rheumatoid arthritis (RA; n=62), mixed connective tissue disease (n=35), dermatomyositis (n=22), polymyositis (n=18), anti-synthetase syndrome (n=3), primary Sjogren's syndrome (n=20), systemic lupus erythematosus (n=12), polymyalgia rheumatica (n=2) and overlap syndrome (n=2); [¶]: p-value for comparison across diagnoses; [+]: IPF (n=499), IPAF (n=244), CTD-ILD (n=248); [§]: Bonferroni-corrected p-values for pairwise comparison between diagnoses for telomere length were as follows: IPF *versus* IPAF (p<0.0001), IPF *versus* CTD-ILD (p<0.0001) and IPAF *versus* CTD-ILD (p=1.0); [f]: restricted to non-Hispanic white patients; [##]: Bonferroni-corrected p-values for pairwise comparison between diagnoses for *MUC5B* MAF were as follows: IPF *versus* IPAF (p=0.00088), IPF *versus* CTD-ILD (p<0.0001) and IPAF *versus* CTD-ILD (p=1.0); [¶¶]: restricted to patients with three or more FVC measurements over a span of ⩾90 days.

Genetic and genomic characteristics

Compared to IPF, age-adjusted LTL was longer for IPAF (−0.17±0.32 *versus* −0.05±0.29; adjusted p<0.0001) and CTD-ILD (−0.04±0.25; adjusted p<0.0001) (table 1). There were twice as many individuals with age-adjusted LTL <10th percentile among those with IPF (31%) than IPAF (16%) or CTD-ILD (13%). Within the CTD-ILD group, rheumatoid arthritis-associated interstitial lung disease (RA-ILD) had shorter age-adjusted LTL (−0.14±0.27) compared to scleroderma-associated interstitial lung disease (SSc-ILD) (−0.02±0.22; adjusted p=0.013) and the other CTD-ILDs (0.00±0.24; adjusted p=0.00042) (table 2). There were more RA-ILD patients with age-adjusted LTL <10th percentile (26%) compared to SSc-ILD patients (12%) and other CTD-ILD patients (6%).

Compared to controls [20], the MAF of the *MUC5B* rs35705950 SNP was higher in IPAF (23.2, 95% CI 18.8–28.2; adjusted p<0.0001) and CTD-ILD (19.9, 95% CI 15.5–25.2; adjusted p<0.0001) patients of non-Hispanic white ethnicity. However, compared to IPF (34.2, 95% CI 31.1–37.5), the *MUC5B* MAF was lower in both IPAF (adjusted p=0.00088) and CTD-ILD (adjusted p<0.0001) patients (table 1). Within the CTD-ILD group, RA-ILD patients of non-Hispanic white ethnicity had a higher *MUC5B* MAF compared to SSc-ILD patients (34.6, 95% CI 24.4–46.3 *versus* 16.6, 95% CI 9.3–26.6; adjusted p=0.040) and other CTD-ILD patients (12.7, 95% CI 7.5–20.4; adjusted p=0.0015) (table 2). In addition, the RA-ILD subgroup had a higher *MUC5B* MAF compared to controls (p<0.0001), while the MAF for the SSc-ILD and other CTD-ILD groups was similar to controls (p=0.19 and p=0.62, respectively). The MAF of the *TOLLIP* rs5743890 SNP was similar across the diagnostic groups and controls [20].

The distribution of LTL and *MUC5B* and *TOLLIP* SNPs between patients with UIP compared to non-UIP pattern were not entirely consistent across diagnostic categories. Telomere length was shorter in the IPAF UIP group *versus* the non-UIP group and there was a higher *MUC5B* MAF in the CTD-ILD UIP group *versus* the non-UIP group (supplemental table S5).

Pulmonary disease progression

Decline in % predicted FVC per year was greater for IPF patients (−5.37, 95% CI −6.10 to −4.66) than IPAF (−1.80, 95% CI −2.70 to −1.00; adjusted p<0.0001) or CTD-ILD patients (−0.64, 95% CI −0.99 to −0.30; adjusted p<0.0001) (table 1). Age-adjusted LTL <10th percentile was associated with a faster decline for IPF and IPAF (figure 1a). For CTD-ILD, the LTL <10th percentile was associated with a trend toward faster decline in % predicted FVC (p=0.028) that did not reach significance (p<0.017) after accounting for multiple testing. The most dramatic difference was in the IPAF cohort where patients with LTL <10th percentile had a −6.43% per year decline compared to −0.86% for those with LTL ⩾10th percentile (p<0.0001). The *MUC5B* or *TOLLIP* genotypes (figures 1b and 1c, respectively) were not associated with change in % predicted FVC per year in IPF, IPAF or CTD-ILD patients.

Patient survival

IPAF patients had longer median transplant-free survival when compared to IPF, but shorter survival when compared to CTD-ILD (table 1). Among the CTD-ILD cohort, the RA-ILD patients had worse transplant-free survival compared to SSc-ILD patients and those with other CTD-ILDs (table 2).

As has been shown previously in other IPF cohorts [11, 15, 16], LTL <10th percentile and the *MUC5B* minor allele were associated with transplant-free survival, but in opposite directions (table 3). For IPAF, shorter LTL (hazard ratio 2.97, 95% CI 1.70–5.20; p=0.00014) and the *MUC5B* minor allele (hazard ratio 1.92, 95% CI 1.18–3.13; p=0.0091) were both associated with worse transplant-free survival. For the CTD-ILD group, the *MUC5B* minor allele was associated with a trend toward worse transplant-free survival (hazard ratio 2.03, 95% CI 1.04–3.95; p=0.038) that did not reach significance (p<0.017) after accounting for multiple testing. The *TOLLIP* genotype was not associated with transplant-free survival in patients with IPAF or CTD-ILD. The results of the overall survival sensitivity analyses were similar (supplemental table S4).

Adding the UIP variable did not change the genomic associations with transplant-free survival. For the IPAF group, LTL <10th percentile (hazard ratio 2.51, 95% CI 1.44–4.39; p=0.0012) and the *MUC5B* minor allele (hazard ratio 1.90, 95% CI 1.12–3.23; p=0.014) were still associated with worse transplant-free survival, while the *TOLLIP* minor allele was not (hazard ratio 0.67, 95% CI 0.35–1.30; p=0.24). In the CTD-ILD group, none of the genomic predictors were associated with transplant-free survival after adding UIP to the model (LTL hazard ratio 1.64, 95% CI 0.80–3.22 (p=0.18); *MUC5B* hazard ratio 1.87, 95% CI 0.89–3.90 (p=0.097); *TOLLIP* hazard ratio 0.67, 95% CI 0.28–1.55 (p=0.35)).

In the model that included LTL and the *MUC5B* genotype as covariates, both were independently associated with transplant-free survival for patients with IPF, but in opposite directions (table 4). For IPAF, LTL <10th percentile was associated with worse transplant-free survival (hazard ratio 2.63, 95% CI 1.47–4.69; p=0.0011) after adjusting for *MUC5B* genotype.

TABLE 2 Characteristics of patients with subtypes of connective tissue disease-associated interstitial lung disease (CTD-ILD)

Characteristic	RA-ILD (n=62)	SSc-ILD (n=74)	Other CTD-ILD[#] (n=112)	p-value[¶]
Age years	60.2±10.5	48.0±11.7	54.1±14.2	<0.0001
Male gender	21 (34)	20 (27)	29 (26)	0.51
Ethnicity				
Non-Hispanic white	40 (65)	41 (55)	59 (53)	0.31
Ever smoker	40 (65)	17 (23)	48 (43)	<0.0001
Family history	4 (6)	1 (1)	1 (1)	0.065
Telomere length				
Observed−expected	−0.14±0.27	−0.02±0.22	0.00±0.24	0.00054[+]
<10th percentile	16 (26)	9 (12)	7 (6)	0.0011[+]
SNPs[§]				
MUC5B rs35705950 MAF	34.6 (95% CI 24.4–46.3)[f] (n=40)	16.2 (95% CI 9.3–26.6)[f] (n=41)	12.7 (95% CI 7.5–20.4)[f] (n=59)	0.00053[##]
TOLLIP rs5743890 MAF	20.5 (95% CI 12.5–31.5)[¶¶] (n=40)	7.7 (95% CI 3.2–16.6)[¶¶] (n=41)	14.4 (95% CI 8.9–22.3)[¶¶] (n=59)	0.072
Disease progression[++]				
ΔFVC % predicted·year^{-1}	−0.59 (95% CI −1.33 to 0.14) (n=89)	−1.03 (95% CI −1.62 to −0.44) (n=54)	−0.41 (95% CI −0.91 to 0.10) (n=89)	0.61
Survival				
Transplant-free survival years (median)	6.32 (95% CI 4.26–NA)	11.88 (95% CI 9.18–NA)	NA (95% CI 9.83–NA)	0.00054

Data are presented as n (%) or mean±SD, unless otherwise stated. RA-ILD: rheumatoid arthritis-associated interstitial lung disease; SSc-ILD: scleroderma-associated interstitial lung disease; SNP: single nucleotide polymorphism; MAF: minor allele frequency; FVC: forced vital capacity; NA: not available. [#]: other CTD-ILD diagnoses include mixed connective tissue disease (n=35), dermatomyositis (n=22), polymyositis (n=18), anti-synthetase syndrome (n=3), Sjogren's syndrome (n=20), systemic lupus erythematosus (n=12), polymyalgia rheumatica (n=2) and overlap syndrome (n=2); [¶]: p-value for comparison across diagnoses; [+]: Bonferroni-corrected p-values for pairwise comparison between diagnoses for telomere length were as follows: RA-ILD *versus* SSc-ILD (p=0.013), RA-ILD *versus* other CTD-ILD (p=0.00042), SSc-ILD *versus* other CTD-ILD (p=1.0) and RA-ILD *versus* non-RA CTD-ILD (p=0.00055); [§]: restricted to non-Hispanic white patients; [f]: comparison of *MUC5B* rs35705950 MAF of non-Hispanic white normal controls (10.7, 95% CI 8.9–12.8) to RA-ILD (p<0.0001), SSc-ILD (p=0.19) and other CTD-ILD (p=0.62); [##]: Bonferroni-corrected p-values for pairwise comparison between diagnoses for *MUC5B* MAF were as follows: RA-ILD *versus* SSc-ILD (p=0.040), RA-ILD *versus* other CTD-ILD (p=0.0015), SSc-ILD *versus* other CTD-ILD (p=1.0) and RA-ILD compared to non-RA CTD-ILD (p=0.00025); [¶¶]: comparison of *TOLLIP* rs5743890 MAF of non-Hispanic white normal controls (14.2, 95% CI 12.1–16.6) to RA-ILD (p=0.18), SSc-ILD (p=0.15) and other CTD-ILD (p=1.0); [++]: restricted to patients with three or more FVC measurements over a span of ⩾90 days.

Discussion

The evaluation of ILD hinges on classification into discrete ILD subtypes to infer expectations regarding disease course, treatment and prognosis. Classification can be challenging when patients do not fit neatly within the IPF and CTD-ILD categories, as is the case for IPAF. In this multicenter cohort study, the clinical characteristics and outcomes of patients with IPAF fall between those of IPF and CTD-ILD. Fewer IPAF and CTD-ILD patients have short LTL (<10th percentile) compared to IPF. However, short LTL is associated with faster lung function decline and worse transplant-free survival in IPAF, similar to IPF. The *MUC5B* MAF is higher in IPAF patients compared to controls and the minor allele is associated with worse transplant-free survival for these patients. The CTD-ILD group as a whole also had higher *MUC5B* MAF compared to controls, but this is largely due to the higher MAF in the RA-ILD subgroup.

Determining if the IPAF classification criteria identifies patients that are truly distinct in terms of disease behaviour, prognosis, or response to therapy compared to IPF or CTD-ILD is clinically important. However, prior studies comparing prognosis of IPAF to either CTD-ILD or IPF demonstrate inconsistent results [7, 21]. Perhaps these inconsistencies are due to differences in cohort composition with regard to LTL and *MUC5B*. In this multicenter cohort study, IPAF patients differ from IPF and CTD-ILD patients in terms of demographics, rate of progression and overall prognosis. In addition, the distribution of the LTL and *MUC5B* genotype differ between IPAF and IPF. Half as many IPAF patients have short LTL compared to IPF, but IPAF patients with short LTL have faster lung function decline and poor survival. In fact, dichotomising IPAF by LTL ⩾10th percentile or <10th percentile distinguishes two groups of patients whose rates of lung function decline resemble those of CTD-ILD and IPF patients, respectively. The *MUC5B* minor allele is overrepresented in patients with IPAF compared to controls, but the minor allele frequency is still significantly lower than in IPF patients. The *MUC5B* minor allele is associated with worse, not better, transplant-free survival in IPAF, which is the opposite of its effect on IPF. Therefore, these genome markers identify specific endotypes within each ILD subgroup that have different rates of progression and survival characteristics.

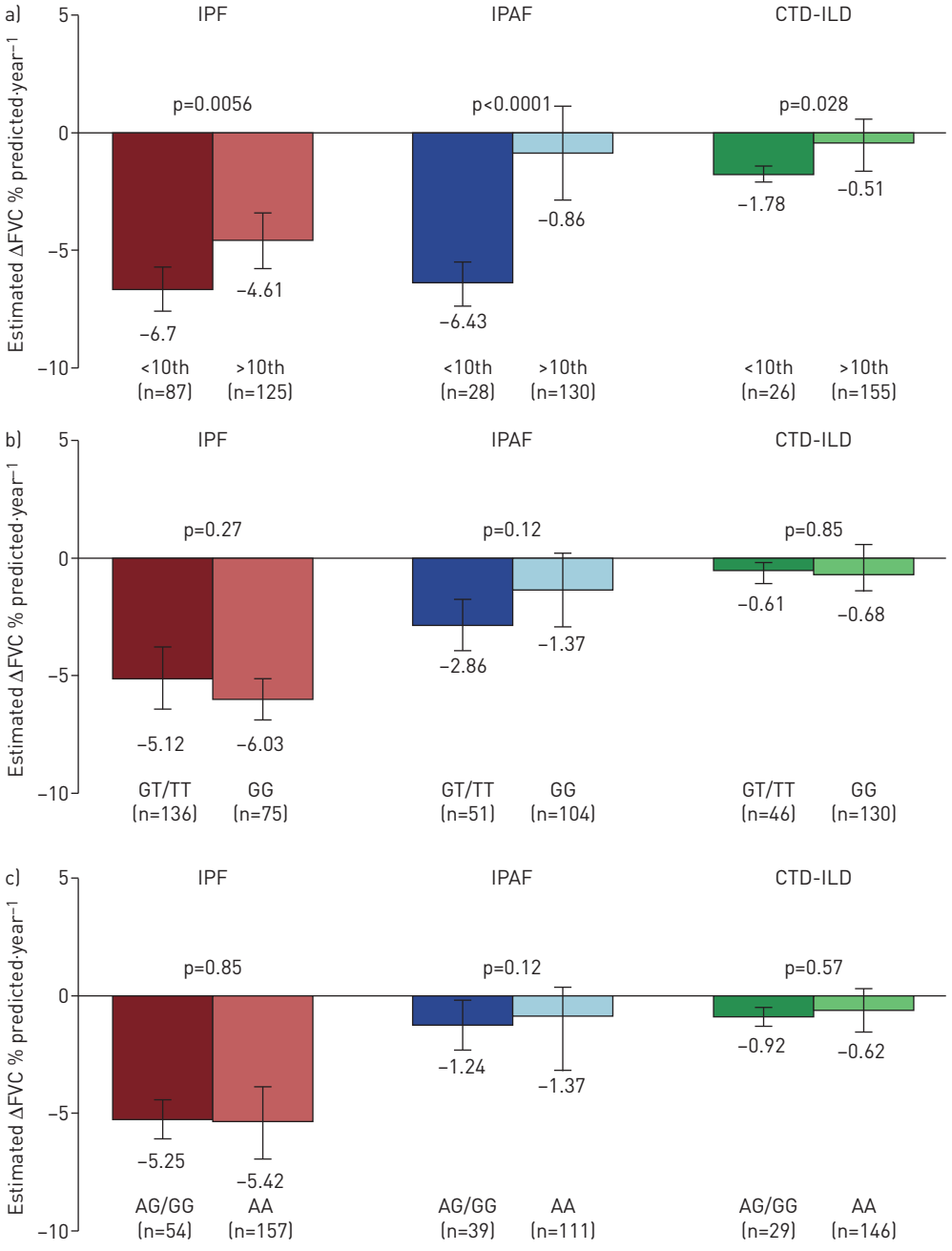

FIGURE 1 Rate of pulmonary disease progression in interstitial lung disease (ILD) patients as measured by the mean change in forced vital capacity (FVC). Estimated change in FVC (% predicted·year^{-1}) for patients with idiopathic pulmonary fibrosis (IPF), interstitial pneumonia with autoimmune features (IPAF) and connective tissue disease-associated interstitial lung disease (CTD-ILD) were stratified as follows: a) by age-adjusted blood leukocyte telomere length (LTL; less than 10th percentile (<10th) or greater than 10th percentile (>10th)); b) by the presence of the *MUC5B* rs35705950 minor allele (GT/TT); c) by the presence of the *TOLLIP* rs5743890 minor allele (AG/GG). This analysis was limited to the subset of patients for which there were at least three spirometry measurements spanning at least 90 days (significant with Bonferroni correction for multiple testing with three predictors (LTL, *MUC5B*, *TOLLIP*) per diagnosis; alpha level of 0.017 per test (0.05/3)).

CTD-ILD represents a collection of various systemic autoimmune disorders that result in lung fibrosis. Patients with CTD-ILD differ from IPF patients in terms of the mechanism of disease, demographics of the affected patients and clinical course. Genomic markers associated with IPF are less prevalent in the CTD-ILD group as a whole. The mean LTL for CTD-ILD patients is only slightly shorter than the expected age-adjusted length [14, 22] and LTL has not been previously associated with transplant-free survival in CTD-ILD patients [14]. In addition, prior studies of patients with SSc-ILD and other CTD-ILDs found no difference in the MAF for *MUC5B* rs35705950 compared to controls [23–26].

TABLE 3 Associations between telomere length and single nucleotide polymorphisms with transplant-free survival for patients with idiopathic pulmonary fibrosis (IPF), interstitial pneumonia with autoimmune features (IPAF) and connective tissue disease-associated interstitial lung disease (CTD-ILD)

Parameter	IPF			IPAF			CTD-ILD		
	n (events)	Hazard ratio (95% CI)	p-value	n (events)	Hazard ratio (95% CI)	p-value	n (events)	Hazard ratio (95% CI)	p-value
Telomere length (<10th percentile)									
Unadjusted	499 (326)	1.92 (1.52–2.44)	**<0.0001**[¶]	244 (102)	2.75 (1.73–4.37)	**<0.0001**[¶]	248 (74)	2.42 (1.3–4.51)	**0.0053**[¶]
Adjusted[#]	386 (232)	1.96 (1.46–2.62)	**<0.0001**[¶]	203 (85)	2.97 (1.70–5.20)	**0.00014**[¶]	197 (52)	1.72 (0.84–3.49)	0.14
***MUC5B* rs35705950 (TT/GT)**									
Unadjusted	495 (324)	0.65 (0.52–0.82)	**0.00018**[¶]	240 (100)	1.52 (1.01–2.28)	0.046	243 (72)	1.92 (1.18–3.12)	**0.0088**[¶]
Adjusted[#]	384 (230)	0.46 (0.34–0.62)	**<0.0001**[¶]	199 (83)	1.92 (1.18–3.13)	**0.0091**[¶]	194 (51)	2.03 (1.04–3.95)	0.038
***TOLLIP* rs5743890 (GG/AG)**									
Unadjusted	495 (324)	1.41 (1.10–1.81)	0.0074	233 (98)	0.65 (0.37–1.13)	0.13	241 (71)	0.90 (0.45–1.83)	0.78
Adjusted[#]	384 (230)	1.32 (0.98–1.79)	0.072	193 (81)	0.57 (0.30–1.08)	0.083	192 (50)	0.72 (0.32–1.66)	0.44

[#]: adjusted for age, gender, non-Hispanic white ethnicity, baseline % predicted forced vital capacity and baseline % predicted diffusing capacity of the lung for carbon monoxide; [¶]: p-values in bold are significant with Bonferroni correction for multiple testing with three predictors (leukocyte telomere length, *MUC5B*, *TOLLIP*) per diagnosis (alpha level of 0.017 per test (0.05/3)).

TABLE 4 Independent associations of telomere length and the *MUC5B* rs35705950 single nucleotide polymorphism (SNP) for transplant-free survival in patients with idiopathic pulmonary fibrosis (IPF), interstitial pneumonia with autoimmune features (IPAF) and connective tissue disease-associated interstitial lung disease (CTD-ILD)

Parameter	IPF (n=384)		IPAF (n=199)		CTD-ILD (n=194)	
	Hazard ratio (95% CI)[#]	p-value	Hazard ratio (95% CI)[#]	p-value	Hazard ratio (95% CI)[#]	p-value
Telomere length (<10th percentile)	2.00 (1.50–2.69)	**<0.0001**[¶]	2.63 (1.47–4.69)	**0.0011**[¶]	1.53 (0.74–3.18)	0.25
***MUC5B* rs35705950 (TT/GT)**	0.45 (0.34–0.61)	**<0.0001**[¶]	1.62 (0.98–2.68)	0.060	1.97 (1.00–3.86)	0.049

[#]: adjusted for telomere length (<10th percentile), *MUC5B* rs35705950 SNP (TT or GT genotype), age, gender, non-Hispanic white ethnicity, baseline % predicted forced vital capacity and baseline % predicted diffusing capacity of the lung for carbon monoxide; [¶]: p-values in bold are significant with Bonferroni correction for multiple testing with three predictors (leukocyte telomere length, *MUC5B*, *TOLLIP*) per diagnosis (alpha level of 0.017 per test (0.05/3)).

Although these genomic markers do not predict clinical outcomes for the combined CTD-ILD group, they may identify a subgroup of CTD-ILD patients (such as those with RA-ILD) who may have a higher risk for poor outcomes. Compared to the CTD-ILD group as a whole, patients with RA-ILD bear a closer resemblance to IPF patients. RA-ILD and IPF patients share demographic features such as older age and a higher proportion of males and smokers [27–30]. In contrast to other CTD-ILDs, patients with RA-ILD often present with radiographic and histopathologic UIP, which is the pathognomonic pattern of fibrosis in IPF [31, 32]. In the current study, not only do RA-ILD and IPF patients have overlapping clinical features, they also have overlapping genomic characteristics. The proportion of RA-ILD patients with LTL <10th percentile is similar to that in IPF patients (25% and 31%, respectively) as opposed to the other non-RA CTD-ILD patients (9%). A recent study by Juge *et al.* [33] found that the *MUC5B* minor allele is overrepresented in patients with RA-ILD and is specifically associated with a UIP pattern. We found that patients with RA-ILD have a similar overrepresentation of the *MUC5B* minor allele as IPF patients (MAF of 34.6 and 34.2, respectively). In contrast, the other non-RA CTD-ILD patients have a similar *MUC5B* MAF to controls (14.4 and 10.7, respectively). A previous study identified rare, likely pathogenic, variants in telomere-related genes (*TERT*, *RTEL1* and *PARN*) in patients with RA-ILD [34] that were similar to those described in sporadic and familial IPF [35–39]. Unfortunately, this study did not provide a large enough sample size to determine if the genomic predictors, namely LTL and the *MUC5B* minor allele, are associated with differential survival risk in RA-ILD as they are in IPF. In particular, it would interesting to see if the *MUC5B* minor allele is associated with worse survival as in IPAF, or better survival as in IPF.

This study has a number of limitations. As an observational cohort study, our results represent associations between the genomic markers and clinical outcomes, and not causal relationships. Genomic DNA was isolated at each site using different methods that may influence multiplex quantitative PCR measurements and biologic samples of fresh blood were unavailable for measurement of telomere length by more precise methods [40]. However, similar trends in telomere length measurements within diagnostic groups are found across sites and the associations between LTL and IPF survival have been replicated by independent investigators using methods of measuring LTL that include flow cytometry, PCR and genomic sequencing [12, 14–16]. Each centre assigned diagnoses based on retrospective review of clinical information and, therefore, availability of testing at each centre may have biased the patient populations. While all IPAF patients fulfilled pre-defined criteria, heterogeneity across sites remained. Unlike IPF, where the accepted diagnostic criteria have been honed over decades, IPAF is a recent designation that will likely undergo revision as the criteria continue to be studied. In our analysis, we attempted to correct for differences by using multivariable models that stratified by cohort. In addition, sample sizes for patients with discrete CTD-ILD subtypes were small, thus limiting our ability to explore the relationship between genomic markers and disease outcomes within CTD-ILD subgroups. Furthermore, we did not assess the influence of treatment on clinical outcomes across genomic characteristics and ILD diagnoses.

This study is the first to characterise the associations between two genomic markers (*MUC5B* SNP and LTL) and clinical outcomes for IPAF and CTD-ILD patients collected from three independent academic medical centres. For patients with IPAF, as with IPF, both of these genomic markers are independently associated with survival. In addition, for IPAF patients, LTL is independently associated with FVC progression. It remains to be seen how these markers might be used in clinical practice and the optimal therapeutic treatment of IPAF patients is not currently clear. Should they be treated with anti-fibrotic medications like IPF patients or immunosuppressive therapies like CTD-ILD patients? Prospective studies

are needed to answer this very important question and to determine if genomic features will identify patients that may have differential responses to specific therapies.

Acknowledgements: We are grateful to all study participants and also to Tyonn Barbera, Cassandra Hamilton and Ross Wilson (University of Texas Southwestern Medical Center, Dallas, TX, USA) for help with patient recruitment and technical excellence.

Conflict of interest: C.A. Newton reports grants from the National Institutes of Health (KL2TR001103), during the conduct of the study. J.M. Oldham reports grants from the National Heart, Lung, and Blood Institute (NHLBI), and personal fees from Genentech and Boehringer Ingelheim, outside the submitted work. B. Ley has nothing to disclose. V. Anand has nothing to disclose. A. Adegunsoye reports speaking fees and honoraria for advisory boards with Boehringer Ingelheim related to idiopathic pulmonary fibrosis (IPF) within the last 12 months. G. Liu has nothing to disclose. K. Batra has nothing to disclose. J. Torrealba reports personal fees for expert panel work from Roche–Ventana and AbbVie, outside the submitted work. J. Kozlitina has nothing to disclose. C. Glazer is a member of a speaker bureau for Genentech. There is no discussion of therapy in this paper. M.E. Strek reports grants and personal fees for advisory board work and speaker fees from Boehringer Ingelheim, and grants from Genentech-Roche, Gilead and MedImmune, outside the submitted work. P.J. Wolters reports grants from Medimmune and Genentech, and personal fees from Roche and Boehringer Ingelheim, outside the submitted work. I. Noth reports personal fees for consultancy from Boehringer Ingelheim, Genentech, Sanofi Aventis and Global Blood Therapeutics, outside the submitted work, and has a patent pending on TOLLIP and IPF pharmacogenomics. C.K. Garcia reports grants from the National Institutes of Health (R01 HL09309) during the conduct of the study and was previously a member of the advisory board for Pliant Therapeutics.

Support statement: The authors acknowledge funding support provided by the National Institutes of Health (awards R01 HL09309 (C.K. Garcia), KL2TR001103 (C.A. Newton), K23HL13890 (J.M. Oldham), KL2TR001870 (B. Ley) and R01 HL130796 (I. Noth)). Funding information for this article has been deposited with the Crossref Funder Registry.

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. *Am J Respir Crit Care Med* 2011; 183: 788–824.
2 Idiopathic Pulmonary Fibrosis Clinical Research Network, Raghu G, Anstrom KJ, *et al.* Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012; 366: 1968–1977.
3 Tashkin DP, Roth MD, Clements PJ, *et al.* Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS II): a randomised controlled, double-blind, parallel group trial. *Lancet Respir Med* 2016; 4: 708–719.
4 Vij R, Strek ME. Diagnosis and treatment of connective tissue disease-associated interstitial lung disease. *Chest* 2013; 143: 814–824.
5 Fischer A, Brown KK, Du Bois RM, *et al.* Mycophenolate mofetil improves lung function in connective tissue disease-associated interstitial lung disease. *J Rheumatol* 2013; 40: 640–646.
6 Fischer A, Antoniou KM, Brown KK, *et al.* An official European Respiratory Society/American Thoracic Society research statement: interstitial pneumonia with autoimmune features. *Eur Respir J* 2015; 46: 976–987.
7 Oldham JM, Adegunsoye A, Valenzi E, *et al.* Characterisation of patients with interstitial pneumonia with autoimmune features. *Eur Respir J* 2016; 47: 1767–1775.
8 Seibold MA, Wise AL, Speer MC, *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011; 364: 1503–1512.
9 Fingerlin TE, Murphy E, Zhang W, *et al.* Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. *Nat Genet* 2013; 45: 613–620.
10 Noth I, Zhang Y, Ma SF, *et al.* Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. *Lancet Respir Med* 2013; 1: 309–317.
11 Peljto AL, Zhang Y, Fingerlin TE, *et al.* Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA* 2013; 309: 2232–2239.
12 Dressen A, Abbas AR, Cabanski C, *et al.* Analysis of protein-altering variants in telomerase genes and their association with MUC5B common variant status in patients with idiopathic pulmonary fibrosis: a candidate gene sequencing study. *Lancet Respir Med* 2018; 6: 603–614.
13 Newton CA, Batra K, Torrealba J, *et al.* Telomere-related lung fibrosis is diagnostically heterogeneous but uniformly progressive. *Eur Respir J* 2016; 48: 1710–1720.
14 Stuart BD, Lee JS, Kozlitina J, *et al.* Effect of telomere length on survival in patients with idiopathic pulmonary fibrosis: an observational cohort study with independent validation. *Lancet Respir Med* 2014; 2: 557–565.
15 Dai J, Cai H, Li H, *et al.* Association between telomere length and survival in patients with idiopathic pulmonary fibrosis. *Respirology* 2015; 20: 947–952.
16 Snetselaar R, van Batenburg AA, van Oosterhout MFM, *et al.* Short telomere length in IPF lung associates with fibrotic lesions and predicts survival. *PLoS One* 2017; 12: e0189467.
17 Ley B, Newton CA, Arnould I, *et al.* The MUC5B promoter polymorphism and telomere length in patients with chronic hypersensitivity pneumonitis: an observational cohort-control study. *Lancet Respir Med* 2017; 5: 639–647.
18 Cronkhite JT, Xing C, Raghu G, *et al.* Telomere shortening in familial and sporadic pulmonary fibrosis. *Am J Respir Crit Care Med* 2008; 178: 729–737.
19 Diaz de Leon A, Cronkhite JT, Katzenstein AL, *et al.* Telomere lengths, pulmonary fibrosis and telomerase (TERT) mutations. *PLoS One* 2010; 5: e10680.
20 1000 Genomes Project Consortium, Auton A, Brooks LD, *et al.* A global reference for human genetic variation. *Nature* 2015; 526: 68–74.
21 Ahmad K, Barba T, Gamondes D, *et al.* Interstitial pneumonia with autoimmune features: clinical, radiologic, and histological characteristics and outcome in a series of 57 patients. *Respir Med* 2017; 123: 56–62.

22 Snetselaar R, van Moorsel CH, Kazemier KM, *et al.* Telomere length in interstitial lung diseases. *Chest* 2015; 148: 1011–1018.
23 Borie R, Crestani B, Dieude P, *et al.* The MUC5B variant is associated with idiopathic pulmonary fibrosis but not with systemic sclerosis interstitial lung disease in the European caucasian population. *PLoS One* 2013; 8: e70621.
24 Stock CJ, Sato H, Fonseca C, *et al.* Mucin 5B promoter polymorphism is associated with idiopathic pulmonary fibrosis but not with development of lung fibrosis in systemic sclerosis or sarcoidosis. *Thorax* 2013; 68: 436–441.
25 Peljto AL, Steele MP, Fingerlin TE, *et al.* The pulmonary fibrosis-associated MUC5B promoter polymorphism does not influence the development of interstitial pneumonia in systemic sclerosis. *Chest* 2012; 142: 1584–1588.
26 van der Vis JJ, Snetselaar R, Kazemier KM, *et al.* Effect of MUC5B promoter polymorphism on disease predisposition and survival in idiopathic interstitial pneumonias. *Respirology* 2016; 21: 712–717.
27 Kelly CA, Saravanan V, Nisar M, *et al.* Rheumatoid arthritis-related interstitial lung disease: associations, prognostic factors and physiological and radiological characteristics–a large multicentre UK study. *Rheumatology (Oxford)* 2014; 53: 1676–1682.
28 Doyle TJ, Dellaripa PF, Batra K, *et al.* Functional impact of a spectrum of interstitial lung abnormalities in rheumatoid arthritis. *Chest* 2014; 146: 41–50.
29 Weyand CM, Schmidt D, Wagner U, *et al.* The influence of sex on the phenotype of rheumatoid arthritis. *Arthritis Rheum* 1998; 41: 817–822.
30 Saag KG, Cerhan JR, Kolluri S, *et al.* Cigarette smoking and rheumatoid arthritis severity. *Ann Rheum Dis* 1997; 56: 463–469.
31 Kim EJ, Collard HR, King TE Jr. Rheumatoid arthritis-associated interstitial lung disease: the relevance of histopathologic and radiographic pattern. *Chest* 2009; 136: 1397–1405.
32 Assayag D, Elicker BM, Urbania TH, *et al.* Rheumatoid arthritis-associated interstitial lung disease: radiologic identification of usual interstitial pneumonia pattern. *Radiology* 2014; 270: 583–588.
33 Juge PA, Lee JS, Ebstein E, *et al.* MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. *N Engl J Med* 2018; 379: 2209–2219.
34 Juge PA, Borie R, Kannengiesser C, *et al.* Shared genetic predisposition in rheumatoid arthritis-interstitial lung disease and familial pulmonary fibrosis. *Eur Respir J* 2017; 49: 1602314.
35 Petrovski S, Todd JL, Durheim MT, *et al.* An exome sequencing study to assess the role of rare genetic variation in pulmonary fibrosis. *Am J Respir Crit Care Med* 2017; 196: 82–93.
36 Tsakiri KD, Cronkhite JT, Kuan PJ, *et al.* Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc Natl Acad Sci USA* 2007; 104: 7552–7557.
37 Armanios MY, Chen JJ, Cogan JD, *et al.* Telomerase mutations in families with idiopathic pulmonary fibrosis. *N Engl J Med* 2007; 356: 1317–1326.
38 Stuart BD, Choi J, Zaidi S, *et al.* Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet* 2015; 47: 512–517.
39 Kannengiesser C, Borie R, Menard C, *et al.* Heterozygous RTEL1 mutations are associated with familial pulmonary fibrosis. *Eur Respir J* 2015; 46: 474–485.
40 Gutierrez-Rodrigues F, Santana-Lemos BA, Scheucher PS, *et al.* Direct comparison of flow-FISH and qPCR as diagnostic tests for telomere length measurement in humans. *PLoS One* 2014; 9: e113747.