




## Basic science

# A genome-wide association study identifying single nucleotide polymorphisms in the *PPFIBP2* gene was predictive for interstitial lung disease in rheumatoid arthritis patients

Shinya Hayashi <sup>1,\*</sup>, Tsukasa Matsubara<sup>2</sup>, Koji Fukuda <sup>2</sup>, Toshihisa Maeda<sup>1</sup>,  
Keiko Funahashi<sup>3</sup>, Marowa Hashimoto<sup>3</sup>, Yoshinori Takashima<sup>1</sup>, Kenichi Kikuchi <sup>1</sup>,  
Masahiro Fujita<sup>1</sup>, Tomoyuki Matsumoto<sup>1</sup>, Ryosuke Kuroda<sup>1</sup>

<sup>1</sup>Department of Orthopaedic Surgery, Kobe University Graduate School of Medicine, Kobe, Japan

<sup>2</sup>Department of Orthopaedic Surgery, Matsubara Mayflower Hospital, Kato, Japan

<sup>3</sup>Research Institute of Joint Diseases, Kobe, Japan

\*Correspondence to: Shinya Hayashi, Department of Orthopaedic Surgery, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe 650-0017, Japan. E-mail: shayashi@med.kobe-u.ac.jp

## Abstract

**Objective:** Genetic polymorphisms might serve as useful prognostic markers for the timely diagnosis of RA. The purpose of this study was to identify genomic factors predictive of the occurrence of interstitial lung disease (ILD) in RA by performing a genome-wide association study of genetic variants, including single nucleotide polymorphisms (SNPs).

**Methods:** The study population included 306 RA patients. All patients were treated with conventional DMARDs, including 6–16 mg MTX per week. Clinical data and venous blood samples were collected from all patients before administration of DMARDs. A total of 278 347 SNPs were analysed to determine their association with ILD occurrence.

**Results:** Several SNPs were strongly associated with ILD occurrence ( $P < 10^{-5}$ ). rs6578890, which is located on chromosome 11 in the intronic region of the gene encoding tyrosine phosphatase receptor type F polypeptide-interacting protein-binding protein 2 (*PPFIBP2*), showed the strongest association with ILD occurrence (odds ratio 4.32,  $P = 10^{-7.93}$ ).

**Conclusion:** *PPFIBP2* could be a useful genetic marker for occurrence of interstitial pneumonia in RA patients and might help to identify the risk of ILD occurrence before RA treatment, thereby improving patient outcomes.

## Lay Summary

### What does this mean for patients?

Interstitial lung disease (ILD), or non-infectious pneumonia, belongs to a class of lung diseases commonly characterized by thickening and scarring of tissue. ILD can lead to morbidity and mortality owing to problems with gas exchange between the lungs and the blood. Genetic polymorphisms, which are variations in DNA sequences between different people, might serve as a predictive marker of disease and help with faster diagnosis. A recent study reported that variations in the *MUC5B* gene were associated with ILD in rheumatoid arthritis patients worldwide. However, the genetic background was different among different populations. We tried to identify predictive markers of ILD in Japanese rheumatoid arthritis patients. We used a genome-wide association study to search for small variations in people's DNA. We found that a variation of the *PPFIBP2* gene had the strongest association with ILD occurrence. Testing for this variation might help doctors to identify ILD occurrence before treatment for rheumatoid arthritis, which could improve patient outcomes.

**Keywords:** genome-wide screening, RA, interstitial lung disease, single nucleotide polymorphisms, *PPFIBP2*

### Key messages

- We analysed a total of 278 347 single nucleotide polymorphisms to determine their association with occurrence of interstitial lung disease in RA.
- rs6578890, encoding *PPFIBP2*, was identified as having the strongest association with occurrence of interstitial lung disease in RA.

Received: 18 May 2022. Accepted: 14 October 2022

© The Author(s) 2022. Published by Oxford University Press on behalf of the British Society for Rheumatology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

## Introduction

Interstitial lung disease (ILD) or non-infectious pneumonia belongs to a class of diffuse lung diseases and represents a group of lung diseases commonly characterized by pulmonary fibrosis or progressive alveolar interstitial sclerosis that can lead to morbidity and mortality owing to respiratory insufficiency [1]. Pulmonary fibrosis is caused by several factors, including exposure to environmental pollutants, such as smoking and inhalation of particulate matter, drug toxicity, radiation exposure, collagen-related vascular diseases and genetics [2, 3]. Smoking is regarded as a major associated risk factor for the development of ILD [1]. A nationwide cohort study demonstrated that smokers had a 1.66-fold increased risk of ILD in comparison to non-smokers [4].

Idiopathic pulmonary fibrosis is the most severe type of ILD. Interestingly, lung transplantation used to be the only available treatment option [1]. Recently, the INBUILD trial demonstrated that nintedanib, an intracellular inhibitor of tyrosine kinases, slowed the rate of decline in forced vital capacity in subjects with progressive fibrosing ILDs [5]. Additionally, the SENSICIS trial demonstrated that nintedanib has antifibrotic and antiinflammatory effects in ILD associated with SSc [6]. These trials suggest that new drug treatments using antifibrotics might be beneficial to treat ILD. Seibold *et al.* [7] reported that a promoter variant in *MUC5B* (rs35705950) was expressed in nearly 50–60% of patients with ILD and was associated with a 6-fold and 20-fold increased risk of ILD for heterozygotes and homozygotes, respectively.

RA is an autoimmune disease characterized by chronic inflammation of the synovial lining of the joints, progressive joint destruction and systemic complications [8, 9]. The prevalence of RA is estimated to range between 0.5 and 1.0% in the general population [8]. As an autoimmune disease, 70–80% of RA patients produce autoantibodies, such as RF and anti-CCP antibodies [10].

Genetic polymorphisms might serve as useful prognostic markers for the timely diagnosis of RA [11]. The International HapMap Project, which has identified variations in the human genome [12], has facilitated the undertaking of genome-wide association studies (GWASs) of genetic variants, including single nucleotide polymorphisms (SNPs) [13]. For example, GWASs have helped in the identification of *PADI4* as a non-MHC genomic locus predictive of RA in the Japanese population [12]. A recent study reported that the *MUC5B* promoter variant rs35705950 is associated with ILD (specifically, with interstitial pneumonia) in RA patients worldwide, including Japanese patients, based on high-resolution chest CT results [14]. The *MUC5B* promoter polymorphism rs35705950 and shorter telomere length are associated with the extent of fibrosis in ILD, while shorter telomere length alone is associated with histopathology findings typical of usual interstitial pneumonia (UIP) and reduced survival in patients with ILD [15]. However, obtaining high-quality imputation results for the *MUC5B* variant has been challenging because of the rare minor allele frequency, which has been estimated at 0.006 in our population-specific imputation reference panel for Japanese patients [16]. Recently, Shirai *et al.* [17] reported that the variant rs12702634 in *RPA3-UMAD1* is associated with a relatively high RA-ILD risk in CT-confirmed fibrosis. However, a follow-up replication study by Higuchi *et al.*

[18] demonstrated no association between rs12702634 and RA-ILD in Japanese patients.

The purpose of this study was to identify genomic factors predictive of ILD occurrence in RA patients by performing a GWAS of genetic variants, including SNPs.

## Methods

### Ethics statement and patient consent

This study complies with the Declaration of Helsinki, and the study protocols were approved by the ethics committee of the Research Institute of Joint Disease, Kobe, and Kobe University Graduate School of Medicine on 11 July 2004. All methods were performed in accordance with the relevant guidelines and regulations of the ethics committees. Informed consent for participation in the study was obtained from the patients before collection of blood samples for genetic analysis.

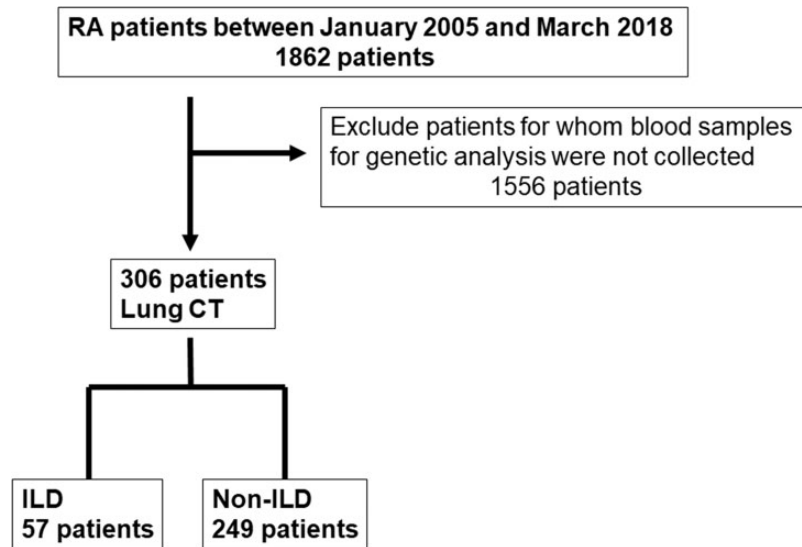
### Patients

This is a retrospective cohort study. Overall, 1862 RA patients who were treated at Matsubara Mayflower Hospital from January 2005 until March 2018 were screened. All patients included in the study met the ACR 1987 revised criteria for RA [19]. Patients for whom blood samples were not collected were excluded from this study. The complete genetic data of 306 patients in the patient registry were analysed. All clinical data and blood samples were collected before any treatment, including treatment with MTX, was started. The patients in this cohort had undergone conventional or high-resolution lung CT, a common practice in the clinic for patients with respiratory symptoms. The average duration of RA was 10.9 years for RA-ILD patients and 10.3 years for non-RA-ILD patients as of March 2018. The ILD status of each subject was determined using CT imaging and classified as UIP, probable UIP or non-specific interstitial pneumonia (NSIP) according to the international criteria [20]. Based on the CT image patterns, UIP, probable UIP and NSIP were defined as RA-ILD (Fig. 1). ILD was diagnosed by two radiologists and one pulmonologist. In the few cases in which the physician's reviews and the radiologist's reports did not agree, the physician re-examined the chest CT images, taking into consideration the radiologist's reports. Besides CT findings, a clinical diagnosis of RA-ILD was also confirmed by referring to medical records. Overall, 57 patients were diagnosed with ILD: 25 patients were assigned in the UIP or probable UIP group and 32 patients in the NSIP group.

All patients were treated with conventional DMARDs, including MTX. MTX doses ranged from 6 to 16 mg/week; dose increases from 6–8 to 16 mg/week followed the Japan College of Rheumatology guideline for the use of MTX in RA patients [21, 22]. Folic acid supplements were administered to all patients treated with MTX; a folic acid tablet was administered the day after MTX treatment. All DNA samples were included in the Japanese GWAS analysis.

### Genome-wide SNP analysis and association study

Samples of 7 ml of venous blood were obtained into glass tubes and kept at 4°C until DNA extraction at Mitsubishi BCL. DNA samples were prepared using the GENTRA PUREGENE DNA isolation kit according to the manufacturer's protocol (catalogue no. 158489; Quiagen, Germany).



**Figure 1.** Study flow chart

DNA integrity was assessed using a Nanodrop 2000 spectrophotometer (Thermo Scientific, USA) to determine the ratio of optical density at 260 nm/280 nm and the concentration of total DNA. Genome-wide SNP genotyping was performed by deCode Genetics (Reykjavik, Iceland) using the Illumina HumanHap300K chip technology (Illumina, San Diego, CA). The genotyping of 317 503 SNPs was performed, including a quality control analysis. SNP genotyping was conducted using Infinium OmniExpressExome-8 Bead Chip Kit (catalogue no. 20024676; Illumina) according to the manufacturer's protocol. SNP genotyping, calling and quality control of samples and SNPs were performed using Illumina GenomeStudio v.2011 software and a cluster file. Genotypes were scored using the GenomeStudio software and GenCall threshold of 0.15. Samples with call rates of >98% were accepted. SNPs were excluded using the following quality control criteria: R mean values of at least one of three clusters <0.25; Cluster Sep values <0.4; or number of no call values >2 on all chromosomes except chromosome Y. After exclusion, 278 347 SNPs were retained for case-control analysis.

### Statistical analysis

The demographic and clinical characteristics of the patients assigned to the ILD (57 patients) and non-ILD (249 patients) groups were compared using the Mann-Whitney *U* test (e.g. age and RA duration), and contingency tables for categorical data were compared using Fisher's exact test (i.e. gender). The frequency of major allelic combinations was compared using Fisher's exact test between ILD and non-ILD patients. Bonferroni's correction was used for multiple comparisons.

## Results

### Patient characteristics

There was no significant difference in disease duration, average stage and class of Steinbrocker classification, and laboratory results, including white blood cell counts, lactate dehydrogenase levels and CRP levels between the ILD and non-ILD groups (Table 1). However, the average age ( $P=0.004$ ), gender (percentage of male patients;  $P<0.001$ ) and average KL-6 (a mucinous glycoprotein expressed on

**Table 1.** Comparison of characteristics and laboratory data between patients with or without interstitial lung disease

Parameter	Patients with interstitial lung disease	Patients without interstitial lung disease	<i>P</i> -value
<b>Patient characteristic</b>			
Number of patients	57	249	
Age, years	66.5	58.6	0.004
Gender, % male	22.6	10.8	<0.001
Duration of RA, years	10.9	10.3	0.737
Steinbrocker stage	3.3	3.3	0.866
Steinbrocker class	2.0	2.1	0.654
<b>Laboratory data</b>			
White blood cell count, $\mu$ l	8647	8091	0.682
Lactate dehydrogenase, U/l	212	198	0.711
CRP, mg/dl	2.84	2.58	0.102
Sialylated carbohydrate antigen KL-6, U/l	443	293	0.037
RF, mg/dl	482	230	0.012
Anti-CCP antibody, U/ml	496	179	0.008

Data are expressed as the mean (s.d.), unless otherwise indicated.

Differences in the continuous demographic variables between ILD and non-ILD patients were analysed using the Mann-Whitney *U* test.

type II pneumocytes;  $P=0.037$ ), RF ( $P=0.012$ ) and CCP antibody ( $P=0.008$ ) titres were significantly different between the two groups (Table 1).

### Genetic association between SNPs and ILD occurrence

Several SNPs strongly correlated with ILD occurrence were identified ( $P<10^{-5}$ ; Table 2). These SNPs included the major homozygous allele SNPs: rs6578890, located on chromosome 11 in the intronic region of *PPFIBP2* ( $P=10^{-7.93}$ ); rs9804373, on chromosome 10 in the intergenic region of *RPA2P2/LOC105378390* ( $P=10^{-5.42}$ ); rs9320420, on chromosome 6 in the intergenic region of *LOC100287612/MARCKS* ( $P=10^{-5.41}$ ); rs2116538, on chromosome 2 in the intronic region of *THSD7B* ( $P=10^{-5.17}$ ); rs2691079, on chromosome 3 in the intergenic region of *GBE1/LOC100289598* ( $P=10^{-5.15}$ ); and rs6758592, on

**Table 2.** Single nucleotide polymorphisms associated with occurrence of interstitial lung disease

Odds ratio	$-\log_{10}P$	Name	Chromosome	Gene symbol	Gene description	Gene location
4.32	7.93	rs6578890	11	PPFBP2	PPFIA binding protein 2	Intron
4.07	5.42	rs9804373	10	RPAP2P2/LOC105378390	Replication protein A2 pseudogene 2/uncharacterized LOC105378390	Intergenic
2.99	5.41	rs9320420	6	LOC100287612/MARCKS	Uncharacterized LOC100287612/myristoylated alanine-rich protein kinase C substrate	Intergenic
2.85	5.17	rs2116538	2	THSD7B	Thrombospondin type 1 domain containing 7B	Intron
3.39	5.15	rs2691079	3	GBE1/LOC100289598	1,4-alpha-glucan branching enzyme 1/uncharacterized LOC100289598	Intergenic
0.32	5.05	rs6758592	2	EPAS1	Endothelial PAS domain protein 1	Intron
2.54	4.99	rs11241716	5	KRT18P16/LINC01170	Keratin 18 pseudogene 16/long intergenic non-protein coding RNA1170	Intergenic
6.34	4.91	rs12063103	1	HIST3H2BB/RNF187	Histone cluster 3, H2bb/ring finger protein 187	Intergenic
6.34	4.91	rs12084511	1	RNF187/LOC100129094	Ring finger protein 187/uncharacterized LOC100129094	Intergenic
2.63	4.81	rs1852598	5	HCN1	Hyperpolarization activated cyclic nucleotide gated potassium channel 1	Intron
5.57	4.78	rs2012593	7	MAGI2	Membrane associated guanylate kinase, WW and PDZ domain containing 2	Intron
3.54	4.77	rs7176074	15	REC114	REC114 meiotic recombination protein	Intron
3.12	4.59	rs1375086	3	GBE1	1,4-Alpha-glucan branching enzyme 1	Intron
2.41	4.52	rs12408079	1	LOC643355/CTTNBP2NL	Uncharacterized LOC643355/CTTNBP2 N-terminal like	Intergenic
10.47	4.44	rs2194829	12	LINC02367	Long intergenic non-protein coding RNA 2367	Intron

Single nucleotide polymorphisms associated with interstitial lung disease (ILD) occurrence are shown. The frequency of major allelic combinations was compared in ILD and non-ILD patients using Fisher's exact test. Bonferroni's correction was used for multiple comparisons. The odds ratios and  $P$ -values ( $-\log_{10}P$ ) of the frequency of major allelic combinations between ILD and non-ILD patients are shown. The chromosome number and the gene symbols, characteristics and location were compared and are also shown.

chromosome 2 in the intronic region of *EPAS1* ( $P = 10^{-5.05}$ ). The odds ratios of the frequency of major allelic combinations between ILD and non-ILD samples is demonstrated in Table 2. The *MUC5B* promoter variant rs35705950 was not significantly correlated with ILD occurrence (Table 2). We also compared the frequency of major allelic combinations between UIP or NSIP and non-ILD patients and found no significant association between genetic variants in these groups.

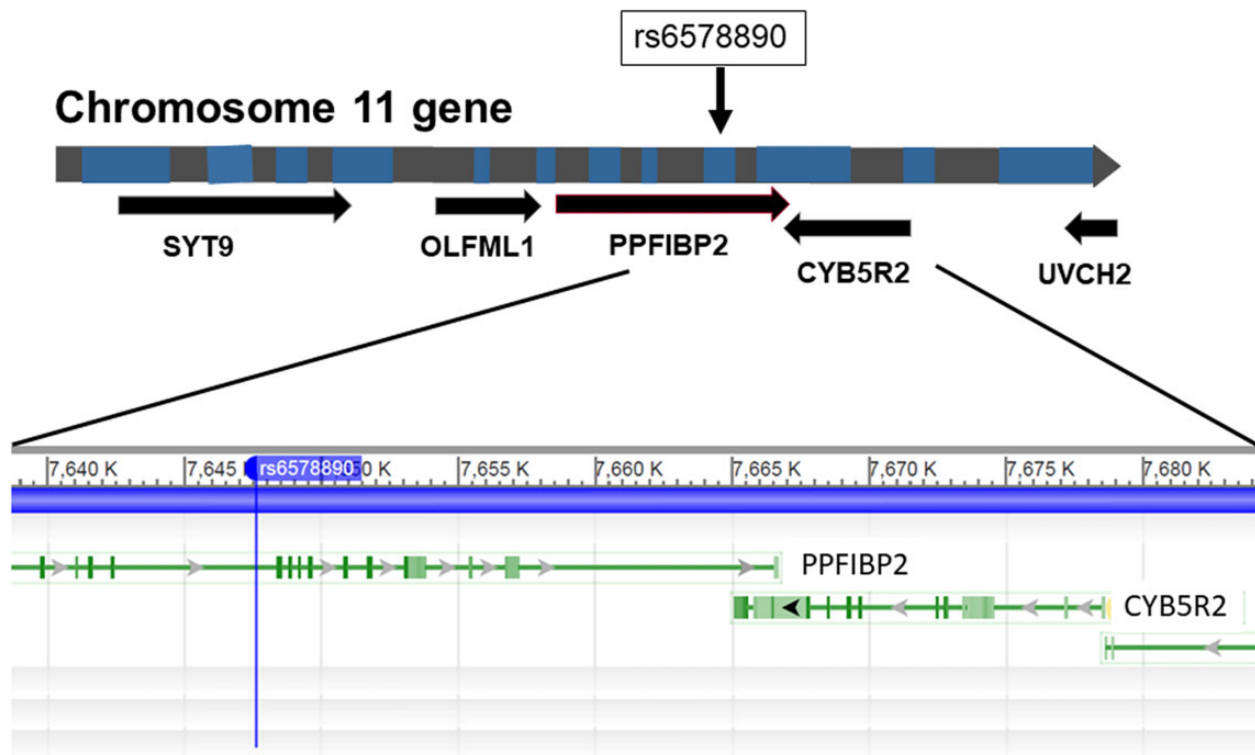
## Discussion

The lifetime risk of ILD in RA patients is ~8%, compared with 1% in the general population [23]. Several risk factors for ILD occurrence have been reported, including smoking, male gender, older age, and high RF and anti-CCP antibody titres [24]. We demonstrated significant differences in average age ( $P = 0.004$ ) and gender ( $P < 0.001$ ) between the ILD and non-ILD groups, and our findings pertaining to these clinical features are comparable to those reported previously. In addition, we showed that the expression levels of RF, anti-CCP antibodies and KL-6 were higher in the ILD group, which is consistent with the results of previous studies [24, 25].

The present study is the first to reveal a significant genome-wide association between rs6578890, an intronic SNP in *PPFIBP2* located on chromosome 11, and ILD occurrence in Japanese RA patients. The protein tyrosine phosphatase receptor type-F polypeptide-interacting protein-binding protein 2, encoded by *PPFIBP2*, belongs to the leukocyte common antigen-related (LAR) protein tyrosine phosphatase-interacting protein (liprin) family [26]. The encoded protein, liprin beta-2, plays a role in axon guidance and neuronal

synapse development by recruiting LAR protein tyrosine phosphatase to the plasma membrane [26]. Liprin beta-2 inhibits breast cancer cell migration through ERK2 signalling [27], and an increased level of mutation in *PPFIBP2* has been reported in fatal cases of prostate cancer [28]. Regarding ILD occurrence, chronic damage of the alveolar epithelium triggers a series of events leading to aberrant tissue repair and destruction of the alveolar structure [29]. The subsequent wound-healing process is driven by a variety of pathogenic events, which have been described in other degenerative diseases and cancer [8]. Further investigation is required to elucidate the function of rs6578890 in aberrant tissue repair and destruction of the alveolar structure in ILD.

Another possible association between rs6578890 SNPs and ILD occurrence should be considered. Figure 2 shows the locus of rs6578890 in *PPFIBP2* that is distal to cytochrome b5 reductase 2 (*CYB5R2*) on chromosome 11. A previous study reported that the SNPs that influence gene expression levels are called expression quantitative trait loci [30]. Chromatin regions, including expression quantitative trait loci, undergo folding to increase their proximity to the genes they regulate, and the folded module affects the expression levels of distant genes via chromatin–chromatin interactions in several diseases, including neurodegenerative diseases [31]. Thus, chromatin, including expression quantitative trait loci, can be folded, and the SNPs modify the expression level of distant genes. In our study, rs6578890 is located distal to *CYB5R2* and can potentially regulate the expression of this gene. Several studies have reported an association between cytochrome members and ILD occurrence [32, 33]. Further experiments are required to elucidate the function of rs6578890 in altering *CYB5R2* expression in ILD patients.



**Figure 2.** Gene map. The locus of rs6578890 in *PPFIBP2* is shown on the gene map from the list of single nucleotide polymorphisms obtained from the NCBI database. rs6578890 is an intronic single nucleotide polymorphisms in *PPFIBP2*, located on chromosome 11 and distal to *CYB5R2*

The present study has a few limitations. First, we did not perform *in vitro* and *in vivo* experiments. Second, the effect of the identified genetic variants on ILD occurrence was not investigated; additional functional analyses are required to elucidate fully the effect of these and other SNPs on the occurrence of ILD. Third, our study comprised a small cohort of patients from one region; therefore, our results might be subject to geographical bias. Further studies with a larger, more diverse patient population should be performed to confirm the observed association.

In conclusion, rs6578890, an intronic SNP located in *PPFIBP2*, might be associated with ILD occurrence in Japanese patients with RA. Currently, the identification of *PPFIBP2* as a genetic predictor of ILD in RA patients can help to determine the risk of ILD occurrence before RA treatment is administered. Understanding the full effect of rs6578890 on the occurrence of RA-ILD will await future studies.

### Data availability statement

The datasets generated and/or analysed during the present study are not publicly available due to the patent in Japan (patent number 4869834, registration date 25 November 2011) but are available from the corresponding author on reasonable request.

### Funding

No funding was received for this study.

**Disclosure statement:** The authors have declared no conflicts of interest.

### References

- Raghu G. Idiopathic pulmonary fibrosis: new evidence and an improved standard of care in 2012. *Lancet* 2012;380:699–701.
- Ryu JH, Colby TV, Hartman TE, Vassallo R. Smoking-related interstitial lung diseases: a concise review. *Eur Respir J* 2001;17:122–32.
- Hubbard R, Venn A, Smith C *et al.* Exposure to commonly prescribed drugs and the etiology of cryptogenic fibrosing alveolitis: a case-control study. *Am J Respir Crit Care Med* 1998;157:743–7.
- Bae W, Lee CH, Lee J *et al.* Impact of smoking on the development of idiopathic pulmonary fibrosis: results from a nationwide population-based cohort study. *Thorax* 2022;77:470–6.
- Flaherty KR, Wells AU, Cottin V *et al.*; INBUILD Trial Investigators. Nintedanib in progressive fibrosing interstitial lung diseases. *N Engl J Med* 2019;381:1718–27.
- Distler O, Gahlemann M, Maher TM. Nintedanib for systemic sclerosis-associated interstitial lung disease. Reply. *N Engl J Med* 2019;381:1596–7.
- Seibold MA, Wise AL, Speer MC *et al.* A common MUC5B promoter polymorphism and pulmonary fibrosis. *N Engl J Med* 2011;364:1503–12.
- Klareskog L, Catrina AI, Paget S. Rheumatoid arthritis. *Lancet* 2009;373:659–72.
- Radner H, Smolen JS, Aletaha D. Comorbidity affects all domains of physical function and quality of life in patients with rheumatoid arthritis. *Rheumatology (Oxford)* 2011;50:381–8.
- Smolen JS, Aletaha D, Barton A *et al.* Rheumatoid arthritis. *Nat Rev Dis Primers* 2018;4:18001.
- Dudbridge F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet* 2013;9:e1003348.
- Suzuki A, Yamada R, Chang X *et al.* Functional haplotypes of *PADI4*, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 2003;34:395–402.
- Ozaki K, Ohnishi Y, Iida A *et al.* Functional SNPs in the lymphotoxin-alpha gene that are associated with susceptibility to myocardial infarction. *Nat Genet* 2002;32:650–4.
- Juge PA, Lee JS, Ebstein E *et al.* MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. *N Engl J Med* 2018;379:2209–19.
- 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;1:309–17.
- Okada Y, Momozawa Y, Sakaue S *et al.* Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. *Nat Commun* 2018;9:1631.
- Shirai Y, Honda S, Ikari K *et al.* Association of the *RPA3-UMAD1* locus with interstitial lung diseases complicated with rheumatoid arthritis in Japanese. *Ann Rheum Dis* 2020;79:1305–9.
- Higuchi T, Oka S, Furukawa H, Shimada K, Tohma S. Lack of association of rs12702634 in *RPA3-UMAD1* with interstitial lung diseases in Japanese rheumatoid arthritis patients. *Biomark Insights* 2022;17:11772719221091758.
- Arnett FC, Edworthy SM, Bloch DA *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
- Raghu G, Remy-Jardin M, Myers JL *et al.*; American Thoracic Society, European Respiratory Society, Japanese Respiratory Society, and Latin American Thoracic Society. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018;198:e44–e68.
- Atsumi T, Yamamoto K, Takeuchi T *et al.* The first double-blind, randomised, parallel-group certolizumab pegol study in methotrexate-naïve early rheumatoid arthritis patients with poor prognostic factors, C-OPERA, shows inhibition of radiographic progression. *Ann Rheum Dis* 2016;75:75–83.
- Tanaka Y, Kuwana M, Fujii T *et al.*; Japan College of Rheumatology guideline for the use of methotrexate in patients with rheumatoid arthritis. 2019 Diagnostic criteria for mixed connective tissue disease (MCTD): From the Japan research committee of the ministry of health, labor, and welfare for systemic autoimmune diseases. *Mod Rheumatol* 2021;31:29–33.
- Bongartz T, Nannini C, Medina-Velasquez YF *et al.* Incidence and mortality of interstitial lung disease in rheumatoid arthritis: a population-based study. *Arthritis Rheum* 2010;62:1583–91.
- Gochuico BR, Avila NA, Chow CK *et al.* Progressive preclinical interstitial lung disease in rheumatoid arthritis. *Arch Intern Med* 2008;168:159–66.
- Kinoshita F, Hamano H, Harada H *et al.* Role of KL-6 in evaluating the disease severity of rheumatoid lung disease: comparison with HRCT. *Respir Med* 2004;98:1131–7.
- Stafford RL, Tang MY, Sawaya MR, Phillips ML, Bowie JU. Crystal structure of the central coiled-coil domain from human liprin-beta2. *Biochemistry* 2011;50:3807–15.
- von Thun A, Birtwistle M, Kalna G *et al.* ERK2 drives tumour cell migration in three-dimensional microenvironments by suppressing expression of Rab17 and liprin-beta2. *J Cell Sci* 2012;125:1465–77.
- Wu Y, Yu H, Zheng SL *et al.* Germline mutations in *PPFIBP2* are associated with lethal prostate cancer. *Prostate* 2018;78:1222–8.
- Gross TJ, Hunninghake GW. Idiopathic pulmonary fibrosis. *N Engl J Med* 2001;345:517–25.

30. Karch CM, Ezerskiy LA, Bertelsen S, Goate AM; Alzheimer's Disease Genetics Consortium (ADGC). Alzheimer's disease risk polymorphisms regulate gene expression in the *ZCWPW1* and the *CELF1* loci. *PLoS One* 2016;11:e0148717.
31. Roussos P, Mitchell AC, Voloudakis G *et al.* A role for non-coding variation in schizophrenia. *Cell Rep* 2014;9:1417–29.
32. Son B, Kwon T, Lee S *et al.* CYP2E1 regulates the development of radiation-induced pulmonary fibrosis via ER stress- and ROS-dependent mechanisms. *Am J Physiol Lung Cell Mol Physiol* 2017; 313:L916–29.
33. Battelli LA, Ghanem MM, Kashon ML *et al.* Crystalline silica is a negative modifier of pulmonary cytochrome P-4501A1 induction. *J Toxicol Environ Health A* 2008;71:521–32.