



HAL
open science

Genotype-Phenotype Relationships in Inheritable Idiopathic Pulmonary Fibrosis: A Greek National Cohort Study

Effrosyni D Manali, Caroline Kannengiesser, Raphael Borie, Ibrahima Ba, Demosthenes Bouros, Aikaterini Markopoulou, Katerina Antoniou, Lykourgos Kolilekas, Andriana I Papaioannou, Vasileios Tzilas, et al.

► **To cite this version:**

Effrosyni D Manali, Caroline Kannengiesser, Raphael Borie, Ibrahima Ba, Demosthenes Bouros, et al.. Genotype-Phenotype Relationships in Inheritable Idiopathic Pulmonary Fibrosis: A Greek National Cohort Study. *Respiration*, 2022, 101 (6), pp.531-543. 10.1159/000520657 . inserm-03798564

HAL Id: inserm-03798564

<https://inserm.hal.science/inserm-03798564v1>

Submitted on 26 Jul 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

GENotype-PHENotype Relationships in inheritable IPF

The GREEK national cohort Study

Effrosyni D. Manali^{1*}, Caroline Kannengiesser^{2,3*}, Raphael Borie^{3,4}, Ibrahima Ba^{2,3}, Demosthenes Bouros^{5,6}, Aikaterini Markopoulou⁷, Katerina Antoniou⁸, Lykourgos Kolilekas⁹, Andriana I Papaioannou¹, Vasileios Tzilas^{5,6}, Argyrios Tzouvelekis¹⁰, Zoe Daniil¹¹, Evangelia Fouka¹², Despoina Papakosta¹², Areti Xyfteri¹³, Anna Karakatsani¹, Stylianos Loukides¹, Ioanna Korbila¹, Ioannis P. Tomos¹, Athanasios K. Konstantinidis¹⁴, Athina Gogali¹⁴, Paschalis Steiropoulos¹⁵, Ilias C. Papanikolaou¹⁶, Chrysa Bazaka¹⁷, Aggeliki Haritou¹⁸, Theodoros Vassilakopoulos¹⁹, Maria Maniati¹, Konstantinos Kagouridis¹, Evangelos Markozannes⁵, Evangelos Bouros⁵, Christina Rampiadou⁷, Georgia Kounti⁷, Athina Trachalaki⁸, Ilias Dimeas¹¹, Theodoros Karampitsakos¹⁰, Panagiotis Lyberopoulos¹, Nikolaos Malamadakis¹, Sofia Spyropoulou²⁰, Patrick Revy²¹, Elodie Lainey²², Philippe Dieudé²³, Khedidja Rebah², Christelle Ménard², Claire Oudin², Cécile Masson²⁴, Aurélie Plessier²⁵, Marie Legendre²⁶, Nadia Nathan²⁷, Aurore Coulomb-L'Hermine²⁸, Annick Clement²⁷, Serge Amselem²⁶, Catherine Boileau², Bruno Crestani^{3,4*}, Spyros Papiris^{1*}.

[The Greek group of ILD investigators]

¹2nd Pulmonary Medicine Department, General University Hospital "Attikon", Medical School, National and Kapodistrian University of Athens, Greece
 fmanali@otenet.gr, i.korbila@gmail.com, papaioannouandriana@gmail.com,
 etomos@gmail.com, loukstel@med.uoa.gr, akarakats@med.uoa.gr,
 konstantinos.kagouridis@gmail.com, liberpit@gmail.com, nmalamadakis@yahoo.gr,
 mar64man@yahoo.com, papiris@otenet.gr

² APHP Hôpital Bichat, Département de Génétique, Paris, France, Université de Paris, Paris, France caroline.kannengiesser@aphp.fr, ibrahima.ba@aphp.fr, christelle.menard@aphp.fr, khedidja.rebah@aphp.fr, claire.oudin@aphp.fr, catherine.boileau@aphp.fr

³ Université de Paris, INSERM UMR 1152, Paris, France

⁴ APHP, Hôpital Bichat, Service de Pneumologie A, FHU APOLLO, Centre de référence des maladies pulmonaires rares, Paris, France ; raphael.borie@aphp.fr, bruno.crestani@aphp.fr

⁵ First Academic Department of Pneumology, Hospital for Diseases of the Chest "Sotiria", Medical School, National and Kapodistrian University of Athens, Athens, Greece. tzilasvasilios@gmail.com, evangelos@me.com, debouros@gmail.com, emarkozannes@yahoo.gr

⁶ Iatriko Medical Center, Athens, Greece, tzilasvasilios@gmail.com, debouros@gmail.com

⁷ Pulmonary Department, "G Papanikolaou" General Hospital, Thessaloniki, Greece, markopoulouk@gmail.com, ram.christina89@gmail.com, kounxat@gmail.com

⁸ Laboratory of Molecular & Cellular Pneumology, Department of Respiratory Medicine, School of Medicine, University of Crete, Greece. kantoniou@med.uoc.gr, atrachalaki@gmail.com

⁹ 7th Pulmonary Department, Athens Chest Hospital "Sotiria", Athens Greece, lykol@yahoo.gr

¹⁰ Department of Respiratory Medicine, General Hospital of Patras, University of Patras, Greece argyrios.tzouvelekis@fleming.gr, thodoriskarampitsakos@gmail.com

¹¹ Department of Respiratory Medicine, Medical School, University of Thessaly, Larissa, Greece, zdaniil@uth.gr, dimel13@hotmail.com

¹² A Department of Pulmonary Medicine, Aristotle University of Thessaloniki, “G. Papanikolaou” Hospital, Exochi, Thessaloniki, Greece evafouka@gmail.com, dpapakos@med.auth.gr

¹³ Messini Messinias Greece aretixyfteri@gmail.com

¹⁴ Department of Pneumology, Medical School, University of Ioannina, Ioannina, Greece, athenagogali@yahoo.com, athkon@yahoo.com

¹⁵ Department of Respiratory Medicine, Medical School, Democritus University of Thrace, University General Hospital Alexandroupolis, Greece, steiropoulos@yahoo.com

¹⁶ Pulmonary Department, Corfu General Hospital, ipapanikolaou76@hotmail.com

¹⁷ Agrinio, Aetolia-Acarmania Greece chrdoct@hotmail.com

¹⁸ Ioannina Greece achdm2@gmail.com

¹⁹ 3rd Department of Critical Care Medicine, Evgenideio Hospital, Medical School, National and Kapodistrian University of Athens, Greece. tvassil@med.uoa.gr

²⁰ Chalkis, Greece spirsofia@yahoo.gr

²¹ INSERM UMR 1163, Laboratory of Genome Dynamics in the Immune System, labellisé Ligue contre le cancer, Université de Paris, Imagine Institute, Paris, France, patrick.revy@inserm.fr

²² Hôpital Robert Debré- APHP, Service d'Hématologie Biologique, Hôpital Robert Debré- APHP, UMRS_1131- Institut Universitaire d'Hématologie, Paris, France, elodie.lainey@aphp.fr

²³ APHP, Hôpital Bichat-Claude Bernard, Department of Rheumatology, INSERM U1152, Université de Paris, Paris, France philippe.dieude@aphp.fr

²⁴ Imagine Institute, Bioinformatics Platform, INSERM UMR 1163, Paris, France. Paris University, Imagine Institute, Université de Paris, Paris, France. cecile.masson@gmail.com, cecile.masson@univ-paris5.fr

²⁵ APHP, Service d'hépatologie, Hôpital Beaujon, Clichy, France, aurelie.plessier@aphp.fr

²⁶ Sorbonne Université, INSERM UMR_S933 Maladies génétiques d'expression pédiatrique, and Assistance Publique Hôpitaux de Paris, U.F. de Génétique moléculaire, Hôpital Armand Trousseau, Paris 75012, France, marie.legendre@aphp.fr, serge.amselem@inserm.fr

²⁷ Sorbonne Université, INSERM UMR_S933, France, Assistance Publique Hôpitaux de Paris, Service de Pneumologie Pédiatrique, Hôpital Armand Trousseau, Centre National de Référence des Maladies Respiratoires Rares RespiRare, Paris 75012, France, nadia.nathan@aphp.fr, annick.clement@aphp.fr

²⁸ Service d'Anatomie et de Cytologie Pathologiques, Sorbonne Université, Hôpital d'Enfants Armand Trousseau, Paris 75012, France, aurore.coulomb@aphp.fr

*These authors contributed equally to this work

Short Title: GEN-PHEN-RE Greek study in inheritable IPF

Corresponding author: Effrosyni D. Manali, Associate Professor, General University Hospital "Attikon", Medical School, National and Kapodistrian University of Athens, Greece, Tel + 302105831184, +302105831163, mobile + 30 6932450858
Email address fmanali@otenet.gr

Number of Tables: 6 tables and 3 at the supplement section

Number of Figures: 2

Word count: Abstract 250 and body text 3813 (4063 in total)

Key words: inheritable pulmonary fibrosis, telomere related genes, surfactant protein related genes, *MUC5B* rs35705950 T risk allele, pathogenic variations, prognosis

Abstract

Background: Monogenic and polygenic inheritance are evidenced for idiopathic pulmonary fibrosis (IPF). Pathogenic variations in surfactant-protein-related genes, telomere-related genes (TRG) and a single-nucleotide polymorphism in the promoter of *MUC5B* gene encoding Mucin 5B (rs35705950 T risk allele) are reported. This French-Greek collaborative study, Gen-Phen-Re-GreekS in inheritable (i)IPF, aimed to investigate genetic components and patients' characteristics in the Greek national IPF cohort with suspected heritability.

Patients and Methods: 150 patients with familial PF, personal-family extra-pulmonary disease suggesting short telomere syndrome and/or young age IPF were analyzed.

Results: *MUC5B* rs35705950 T risk allele was detected in 103 patients (90 heterozygous, 13 homozygous, allelic frequency of 39%), monoallelic TRG pathogenic variations in 19 patients (8 *TERT*, 5 *TERC*, 2 *RTEL1*, 2 *PARN*, 1 *NOP10* and 1 *NHP2*) and biallelic *ABCA3* pathogenic variations in 3. Overlapping *MUC5B* rs35705950 T risk allele and TRG pathogenic variations were shown in 11 patients (5 *TERT*, 3 *TERC*, 1 *PARN*, 1 *NOP10*, 1 *NHP2*), *MUC5B* rs35705950 T risk allele and biallelic *ABCA3* pathogenic variations in 2. In 38 patients neither *MUC5B* rs35705950 T risk allele nor TRG pathogenic variations were detectable. Kaplan-Meier curves showed differences in time-to-death ($p=0.025$) where patients with *MUC5B* rs35705950 T risk allele alone or in combination with TRG pathogenic variations presented better prognosis.

Conclusion: The Gen-Phen-Re-GreekS in iIPF identified multiple and overlapping genetic components including the rarest, underlying disease's genetic "richesse",

complexity and heterogeneity. Time-to-death differences may relate to diverse IPF pathogenetic mechanisms implicating “personalized” medical care driven by genotypes in the near future.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic and progressive fibrotic interstitial lung disease (ILD) of unknown pathogenesis, leading to irreversible respiratory failure and death [1]. No curative treatment is available. IPF was until recently considered a sporadic disease of the elderly, limited to the lungs [1]. This traditional approach has been challenged by the observation that PF may also be encountered in children and young adults, that it develops 10 times more often in members of families with an IPF patient, and that it may be a feature of a multisystem, inherited disease named telomeropathy where it coexists with bone marrow failure and/or liver disease [2, 3, 4]. The above refer to “heritability” and proved to be related mainly to pathogenic variations in telomere-related genes (TRG) or surfactant-protein-related genes (SPRGs) in monogenic inheritance and a single-nucleotide polymorphism in the promoter of the *MUC5B* gene encoding Mucin 5B (rs35705950 T risk allele) in polygenic inheritance [5, 6, 7, 8, 9]. The implicated genes influence biology of two distinct types of lung epithelial cells: the type 2 alveolar epithelial cell (AEC2) and the bronchial secretory cell. SPRG pathogenic variations disrupt AEC2 homeostasis by induction of endoplasmic reticulum stress (ER) leading to a significant increase in the expression of mesenchymal markers and profibrotic mediators and local dysregulation of immune responses [9]. TRG pathogenic variations are related to loss of telomerase complex activity or stability permitting telomere shortening and inducing activation of DNA-damage mechanisms, premature senescence and profibrotic remodeling. TRG pathogenic variations are also associated with the dysregulation of cells in tissues such as the bone marrow and the liver leading to “short-telomere syndrome” (STS) [4]. *MUC5B* rs35705950 T risk allele carriership has been found to increase the expression of the mucin5B in the distal lung causing

mucociliary dysfunction and ER stress triggering thus a vicious cycle of injury/repair/regeneration at the bronchoalveolar junction leading to fibrosis [9, 10].

The SPRGs which pathogenic variations predispose to the development of PF include those encoding surfactant proteins A1 (*SFTPA1*), A2 (*SFTPA2*), C (*SFTPC*), the ATP-binding cassette-type family A member 3 transporter (*ABCA3*) and the NK2 homeobox 1 (*NKX2-1*) transcription factor [5, 11, 12, 13]. The TRG pathogenic variations, major monogenic cause of PF, included initially only heterozygous mutations in telomerase reverse transcriptase (*TERT*) and the RNA component of the telomerase complex (*TERC*), but in a few years embraced more than eight new genes mutations: *RTEL1*, *PARN*, *NAF1*, *DKC1*, *TINF2*, *NOP10*, *NHP2* and *ZCCHC8* [14, 15, 16, 17, 18]. Finally, *MUC5B* rs35705950 T risk allele carriership is the strongest polygenic risk factor for the development of IPF in the overall population [8].

Known genetic causes may explain circa 30% of inheritable (i)PF cases as report studies from the USA and Northern Europe [19, 20]. Few data exist regarding the rest of the world, although scientific evidence underscores the genetic complexity and heterogeneity of iIPF [9]. This French-Greek collaborative study, Gen-Phen-Re-GreekS in (i)IPF, aimed to investigate genetic components in the Greek national IPF cohort with suspected heritability and define patients' characteristics.

Methods

This is a retrospective, observational, non-interventional, collaborative French-Greek study including all patients from the Greek national cohort with suspected iIPF. The study took place from January 2015 to January 2020 and patients were included if they presented at least one of the following: a) familial PF (FPF), that is a family

history of two or more relatives with fibrotic ILD (f)-ILD, b) IPF at a younger age (<55 years old), c) personal/-family extra-pulmonary disease suggesting STS such as a history of bone marrow failure, thrombocytopenia or myelodysplasia, early hair graying or cryptogenic liver cirrhosis [19]. Classic cutaneous triad reported in dyskeratosis congenita (DC) was not present in our patients. Patients were referred for evaluation from specialized referral centers and dedicated private practices from all over Greece to the 2nd Pulmonary Medicine Department, General University Hospital “Attikon” (Athens, Greece). Genetic analysis was performed in France, specifically at the genetic laboratory of Bichat Hospital (Paris, France), for TRG and *MUC5B* rs35705950 T risk allele testing, and at the molecular genetics’ laboratory of Armand Trousseau Hospital (Paris, France) for SPRGs testing. Both are reference academic facilities belonging to the Respifil network including adult and pediatric centers with a special expertise in rare lung diseases (www.respifil.fr). Lung fibrosis diagnosis was based on a multidisciplinary team discussion (MDD) including clinical data, lung high resolution computed tomography (HRCT) and pathological evaluation according to ATS/ERS/JRS/ALAT official guidelines [1]. Clinical information was collected in a standardized manner (See online Patients and Methods Supplement). Treating physicians performed clinical work-up and management. Pulmonary function tests (PFTs) were recorded at diagnosis. Death data were recorded until January 15th 2020, except for cases lost to follow-up. Seven patients of the cohort have been already reported in the literature [16, 21, 22]. All patients signed written informed consent and the study was approved by the Medical Ethics Committee of General University Hospital “Attikon”, Athens, Greece (EBA 585/13-1-15).

Genetic analysis

Molecular explorations were performed by Sanger sequencing, capture-based next generation sequencing panels (Haloplex / Agilent or Nimblegen / Roche) or whole exome sequencing (WES), as previously described [12, 19]. Germline variants were interpreted according to the American College of Genetics and Genomics guidelines and the European Society for Human Genetics recommendations [23]. Only variants classified as clearly pathogenic (class 4 or 5) or suspected to be pathogenic as defined in [19] by VUSD (variant of unknown significance with a working diagnosis of damaging) are reported in this manuscript. A few results were also discussed in MDD dedicated to suspected monogenic pulmonary fibrosis [24]. Most of the patients recruited for family history, relatively young age (<55 years old) or extra pulmonary signs were explored for telomere disease (N=148), excluding very young patients (< 20 years old) suspected to surfactant disease and revealed to be carrier of a mutation in surfactant pathway (N=2). A total of 9 patients were tested for surfactant pathway mutations due to early onset pulmonary fibrosis presentation (age <40 years old) or associated personal or familial lung cancer.

Statistical analysis

Categorical variables are presented as n (%), whereas numerical variables are presented as mean±standard deviation (SD) or median (interquartile ranges) for normally distributed and skewed data, respectively. Normality of distributions was checked with Kolmogorov-Smirnov test. Comparisons between patients with and without pathogenic mutations were performed using chi-square tests for categorical data, as well as unpaired t-tests and one-way ANOVA or Mann-Whitney U-tests and Kruskal–Wallis *H* test for normally distributed or skewed numerical data, respectively. For analyzing the specific sample pairs for stochastic dominance, Dunn's test was used. Cox regression univariate and multivariate analyses were performed in

order to evaluate the influence of factors known to be associated with outcomes in mortality. Kaplan-Meier survival curves were also performed in order to estimate the difference of time to death in months between different genotypes. Statistical significance was calculated by log rank test. Data were analyzed using SPSS 18.0 for Windows (SPSS Inc, Chicago, IL, USA) and Graph Pad prism 6 (La Jolla CA, USA). p-values <0.05 were considered statistically significant.

Results

Patients

During the study period we identified 150 probands from 150 non-related families from all over Greece presenting with iIPF (Table 1). Only one proband per family was included in the analysis. Males were 73% with a median (IQR) age-at-diagnosis of 67.5 (55.0-74.2) year-old, 62.7% ever-smokers, 75% with a definite/probable UIP or CPFE pattern on HRCT of the chest. Histology was available in 26 (17.3%) patients. Presenting symptoms were either cough, or dyspnea or both in 93% of patients and fibrogenic exposure could be identified in 57 (38%). Median values (IQR) for FVC%pred, TLC%pred and DLCO%pred upon inclusion in the study were 75.5 (59.9-93.2), 65 (52.4-77.8) and 48.8 (36.0-62.0) respectively. Combined estimation of Gender, Age, Predicted FVC and DLCO (GAP) score for IPF mortality was median (IQR) value of 2 (1.0, 2.0) for GAP stage. Median values of hematological parameters were within normal limits.

Genetics

Genetic analyses were performed due to FPF in 95 (63.3%) patients, for young age in 23 (15.3%), and personal and/or family extra-pulmonary disease suggestive of STS in 22 (14.7%) and 10 (6.7%) patients respectively; 32 (21.3%) patients were tested for

more than 1 indication (Table 1 supplement). Median (IQR) numbers of affected members in a family and of generations with affected members were 2 (2.0, 3) and 1 (1.0, 2) respectively. Parameters suggestive of personal or familial STS such as anemia, macrocytosis, thrombocytopenia and liver cirrhosis were encountered in 24.3%, 12.5%, 14% and 18.7% of patients respectively. Further characteristics related to what led to suspicion of a mutation such as the number of affected members in a family, the number of generations with affected members and parameters suggestive of STS are also shown in Table 1 supplement. Allele frequency for *MUC5B* rs35705950 T risk allele was 39% corresponding to 103 patients (90 in heterozygous and 13 in homozygous state); Monoallelic TRG pathogenic variations were detected in 19 patients (8 *TERT*, 5 *TERC*, 2 *RTEL1*, 2 *PARN*, 1 *NOP10* and 1 *NHP2*) (Table 2 supplement); surfactant protein-related genes pathogenic variations were detected in 3 patients (biallelic *ABCA3* rare pathogenic variations). *SFTPA1* or *SFTPA2* rare variants were not detected in those patients tested. Combined genotypes such as co-existent *MUC5B* rs35705950 T risk allele with TRG pathogenic variations (5 *TERT*, 3 *TERC*, 1 *PARN*, 1 *NOP10*, 1 *NHP2*) were shown in 11 patients, *MUC5B* rs35705950 T risk allele and biallelic *ABCA3* pathogenic mutations in 2 patients. In 38 patients neither *MUC5B* rs35705950 T risk allele nor TRG pathogenic mutations was detectable (Table 2) (Figure 1).

Factors associated with outcome of the entire cohort

Median (IQR) age at death for all was 75 (64.2, 82.0) years. Median (IQR) survival of the cohort was 3 (1, 6) years. Forty-two patients (28%) deceased during the study period. In univariate analysis, older age at diagnosis, higher GAP index and stage, lower values of FVC%pred, TLC%pred, and DLCO%pred, higher FEV₁/FVCx100 at diagnosis, increased monocyte count and thrombocytopenia were associated with an

increased risk of death whereas *MUC5B* rs35705950 T risk allele carriership and antifibrotic treatment with a decreased risk of death (Table 3). In multivariate analysis remained independent predictors of mortality associated significantly with increased risk of death the higher age at diagnosis, lower values of FVC%pred, higher values of FEV₁/FVCx100 at diagnosis, presence of thrombocytopenia and increased monocytes count (Table 3).

Distribution of HRCT and histology patterns among the genotypes

In 90 patients (60.0%) with *MUC5B* rs35705950 T risk allele (either homozygous or heterozygous) without a monogenic cause presented on HRCT definite UIP 46 patients (51.1%) and probable UIP 15 patients (16.7%) (Table 4). Histopathology data of lung biopsy were available in 12 patients (13.3%) and UIP observable in 9 patients (75%). In 8 patients (5.3%) with TRG pathogenic variations alone, the prevalent HRCT pattern was indeterminate and observed in 3 patients (37.5%). Histopathology data were available in 4 patients (50%) and UIP, or UIP mixed was observed in 3 (75%). In 11 patients (7.3%) with combined TRG pathogenic variations and *MUC5B* rs35705950 T risk allele, the HRCT pattern was definite/probable UIP alone or combined emphysema. Histopathology data in 4 patients (36.4%) confirmed UIP. In 38 patients (25.3%) with neither *MUC5B* rs35705950 T risk allele nor TRG pathogenic variations, all HRCT patterns were observed with definite UIP shown in 16 patients (42.1%). Histopathology features available for 4 patients (10.5%) confirmed UIP in 2 patients and NSIP in 1 patient (and 1 missing report). At post hoc analysis, carriers of TRG pathogenic variations alone were younger at diagnosis in comparison to *MUC5B* rs35705950 T risk allele carriers alone and were submitted more frequently to lung biopsy (50%) (Table 4). Finally, regarding the 3 patients (2%) with *ABCA3* pathogenic variations, HRCT revealed only indeterminate or alternative

ILD patterns and histopathology available in 2 patients showed overlapping features of nonspecific interstitial pneumonia (NSIP), desquamative interstitial pneumonia (DIP) and organizing pneumonia (OP) (Table 3 supplement).

Distribution of patients' main parameters among genotypes

Classifying patients according to their genetic background, no significant differences were detected concerning fibrogenic exposure, symptoms, functional status, extrapulmonary manifestations such as anemia, thrombocytopenia and liver disease, GAP scores and antifibrotic treatment (Table 5). Upon initial evaluation of patients, younger age at diagnosis, coexistence of multiple traits of heritability including macrocytosis and multigenerational families favors a genetic predisposition related to TRG pathogenic variations (Table 6).

Survival between genotypes

Kaplan-Meier curves revealed a significant difference to the time to death between genotypes ($p=0.025$). Patients with TRG pathogenic variations alone or with neither *MUC5B* rs35705950 T risk allele nor TRG pathogenic variations presented the worst prognosis while those with *MUC5B* rs35705950 T risk allele either with or without TRG pathogenic variation presented better prognosis (Figure 2).

Discussion

This French-Greek collaborative study, Gen-Phen-Re-GreekS in iIPF, identified multiple and overlapping genetic components including extremely rare ones among patients with iIPF in the Greek national cohort. In detail frequency for *MUC5B* rs35705950 T risk allele was 39% corresponding to 103 patients (90 heterozygous, 13 homozygous), monoallelic TRG pathogenic variations were detected in 19 patients (8

TERT, 5 *TERC*, 2 *RTEL1*, 2 *PARN*, 1 *NOP10* and 1 *NHP2*) and biallelic *ABCA3* pathogenic variations in 3. Overlapping *MUC5B* rs35705950 T risk allele and TRG pathogenic variations (5 *TERT*, 3 *TERC*, 1 *PARN*, 1 *NOP10*, 1 *NHP2*) were shown in 11 patients, and *MUC5B* rs35705950 T risk allele and biallelic *ABCA3* pathogenic variations in 2. In 38 patients neither *MUC5B* rs35705950 T risk allele nor TRG pathogenic variations were detectable.

Every patient was genotyped for *MUC5B* rs35705950 T risk allele. In 90 patients, the *MUC5B* rs35705950 T risk allele was found alone, in 13 in combination with monogenic pathogenic variations. The reported frequency for *MUC5B* rs35705950 T risk allele of 39% is higher than the 3.8% in historic Greek controls, 3.4% of Asian IPF populations but comparable to the 25% to 40% reported in North European, Mexican and the USA IPF cohorts [8, 19, 25, 26, 27, 28, 29]. Carriers of *MUC5B* rs35705950 T risk allele alone were found to have strong familial traits like 2 or more affected members in family in 81% of cases, multigenerational affected families in 40% and signs of STS in 15%. Phenotypically in comparison to the other genotypes, they were characterized by female gender prevalence, older age at diagnosis and UIP imaging and histology patterns. Our findings are in agreement with both the studies of Seibold M and co-workers and van der Vis JJ and co-workers who recognized *MUC5B* rs35705950 T risk allele as the strongest genetic risk factor for the development of FPF [8, 27] and confirm the observations of Chung JH and co-workers and of Juge PA and coworkers that *MUC5B* rs35705950 T risk allele carriership is related to UIP HRCT pattern [25, 26]. All the above support the elaboration of distinctive *MUC5B* rs35705950 T risk allele driven phenotypic characteristics in iPF. Interactions with other unidentifiable genetic or environmental

factors should be further evaluated although the present study did not reveal significant influence of smoking and fibrogenic exposures.

TRG pathogenic variations carriers represented 13% of the Greek cohort and included classical *TERT* and *TERC* mutations, rare *PARN* and *RTEL1*, ultra-rare *NHP2* mutation and the first ever reported heterozygous *NOP10* mutation in a context of FPF. TRG pathogenic variations are implicated in up to 20% of FPF cases. There are numerous components in the telomerase complex including *TERT*, *TERC*, *PARN*, *RTEL1*, *NAF1*, *DKC1* and *TINF2* which are essential for normal function and new genetic variants are continually being identified. *NOP10* along with *DKC1*, *NHP2* and *GAR1* are essential for *TERC* stability and telomere maintenance. Our observation that the *NOP10* firstly described mutation and *NHP2* are novel causes of FPF and STS amplify the role of TRG pathogenic variations in iPF highlighting some unique aspects of telomere biology [16, 30]. Patients were examined divided in those with TRG pathogenic variations alone (5.3%) and those with the combination of TRG pathogenic variations and *MUC5B* rs35705950 T risk allele (7.4%). Patients with TRG pathogenic variations alone were characterized by younger age at diagnosis (55 years old) and multiple signs of heritability such as personal/-familial STS and multigenerational families. All these features remained identifiable in the TRG pathogenic variation and *MUC5B* rs35705950 T risk allele combination group, except from the age at diagnosis that was much older. In lung biopsy, the TRG pathogenic variation alone group presented NSIP features in one quarter of patients whereas the combination group was characterized by UIP only. Our findings confirm that TRG pathogenic variations are the major monogenic cause in suspected IPF and reveal that the combination with *MUC5B* rs35705950 T risk allele in 58% of patients leads to phenotypic changes that could be attributed to additional effects. In this study very

rare variants in *TERC*, *PARN*, *RTEL1*, *NOPI0* and *NHP2* predominate *TERT* ones underlying the originality of our population. Phenotypic characteristics of the Greek patients on the other hand were very similar to the French and American cohorts including younger age at diagnosis and STS [19, 20]. The limited number of rare TRG pathogenic variations precludes further comparisons [15, 16, 17, 18]. In the only combined analysis of rare and common variants in 1510 patients with sporadic IPF, Dressen A and co-workers demonstrated that 3% of carriers of *MUC5B* rs35705950 T risk allele carried a rare variant of *TERT* compared to 7% in the non-risk allele group, speculating that the presence of TRG pathogenic variations excludes that of *MUC5B* rs35705950 T risk allele [31]; not the case from the Greek cohort. Dressen and co-workers retained variants restricted to exonic SNPs with a minor allele frequency under 1% in the general population. Missense Variants included in the rare variant analysis were high impact variants (e.g stop gain, stop loss, frameshift) or missense variants with a PolyPhen score predicted as damaging (≥ 0.5). The variants retained as pathogenic in our study are more rare (allelic frequency = 0.001).

Three patients out of 9 that were tested represented the genotype group of SPRGs mutations. Biallelic *ABCA3* pathogenic variations were exclusively identified whereas monoallelic *SFTPA1* or *SFTPA2* rare variants were not detected probably due to the limited number of patients tested. Besides the very small number of patients, distinctive characteristics were recognized such as early onset ILD, in childhood, in early adulthood and at 35 years old. All presented finger clubbing, low DLCO in the absence of smoking or fibrogenic exposures. Indeterminate or alternate HRCT patterns corresponded to biopsy-proven NSIP, OP and DIP. The patient with a history of FPF presented also homozygous for *MUC5C* rs35705950 T risk allele. During the study period, no deaths were reported. This study underlines the genotype-phenotype

diversity associated with biallelic *ABCA3* pathogenic variations [13, 32, 33, 34] and shows that biallelic *ABCA3* pathogenic variation related ILD patients surviving from childhood to adulthood are extremely rare but do exist especially when mild mutations allowing a more progressive and benign course are identified [22]. The potential clinical impact of combination with *MUC5B* rs35705950 T risk allele remains to be defined [35].

In one quarter of our patients, no identifiable monogenic or polygenic genetic component was found eliciting a lot of questions about underlying inheritable biology. Those patients were diagnosed at an older age, had the highest rate of smoking and fibrogenic exposures compared to the other genotypes, presented all the range of HRCT and histology patterns with a modest predilection for UIP and had less often multigenerational families. Exome or genome sequencing could increase the possibility to identify novel genes helping to overcome the challenges of unexplained heritability [9]. In that perspective the detailed phenotypic characterization of patients as undertaken in the present study could be of significant help.

The median survival of our cohort was 3 years; very similar to the survival reported for the American cohorts with TRG pathogenic variations related PF [20, 36]. In the French cohort, the survival was found much longer (7.3 years) but as the authors state this could be attributed to the fact that half of the patients presented ILD diagnoses with better prognosis other than IPF [19]. The evaluation of predictors of mortality identified novel independent parameters such as the monocytes count indicating that the ominous impact of increased monocytes in patients with sporadic IPF associated to their differentiation to pro-fibrotic macrophages and fibroblasts could also apply to inheritable disease [37].

A significant difference in the time-to-death was shown in Kaplan-Meier analysis. Patients with TRG pathogenic variations alone or patients with neither *MUC5B* rs35705950 T risk allele nor TRG pathogenic variations presenting worst prognosis. The observed differences in the *MUC5B* rs35705950 T risk allele alone and the combination TRG pathogenic variations groups could be explained by the effect of the *MUC5B* rs35705950 T risk allele on survival based on the existing but recently disputed evidence that this variant associates in an allele-dose dependent manner with better survival in IPF patients, especially in those with familial traits [26, 27, 38, 39]. The dismal prognosis of the TRG pathogenic variations carriers on the other hand is already acknowledged and bears significant clinical implications regarding both transplantation and treatment decisions [39].

Among the limitations of our study are the absence of central reading of HRCT scans and histology. However, all patients were managed in specialized referral ILD centers minimizing the risk of significant inaccuracies. The measurement of telomere length, for obvious geographical reasons, was technically impossible in every patient with TRG pathogenic variations. However, in older patients with lung disease the shortening is less clear [4], suggesting that the measurement in our unexplained older IPF patients would not have been so helpful. Moreover, SRPGs pathogenic variations testing was selectively performed in patients with young age ILD slightly underestimating its true prevalence. Finally, other low risk alleles associated to the development of IPF have not been tested in the present study [40].

To the best of our knowledge, this is the first cohort study that defines characteristics of patients with iIPF related to all the range of major identifiable genetic components and attempts an analysis of patients with both isolated and combined genetic traits. It describes classical, rare, ultra-rare and novel TRG pathogenic variations broadening

the evidence about the importance of impaired telomere maintenance pathway in the pathogenesis of monogenic inheritable pulmonary fibrosis and provides evidence about rarely encountered patients with *ABCA3* pathogenic variations who survive from childhood to adulthood. The role of *MUC5B* rs35705950 T risk allele found in the majority of patients is explored regarding the development of iPF and the elaboration of *MUC5B* rs35705950 T risk allele driven phenotypic characteristics, when this allele is alone or combined with a monogenic cause.

In conclusion, the Gen-Phen-Re-GreekS in iIPF identified multiple and overlapping genetic components including the rarest, underlying disease's genetic "richesse", complexity and heterogeneity. Time-to-death differences may relate to diverse IPF pathogenetic mechanisms implicating "personalized" medical care driven by genotypes in the near future.

Statement of Ethics

This study protocol was reviewed and approved by the Medical Ethics Committee of General University Hospital "Attikon", Athens, Greece (EBA 585/13-1-15). All patients signed written informed consent to participate in the study.

Conflict of interest Any conflict of interest related to this work is reported in detail in the conflict-of-interest form of each author submitted along with the manuscript

Funding Sources This work (exome sequencing) was supported by a grant "FPI-SPC" from Université Sorbonne Paris Cité (B. C.) and the Chancellerie des Universités de Paris (legs Poix; C. K.).

Author's contributions: EDM and SAP had major contribution to the concept and design of the study, to the acquisition, analysis and interpretation of data, had access to all data, supervised the accuracy and integrity of any part of the work and wrote the manuscript; CK had major contribution to the concept and design of the study, performed the genetic analysis regarding the telomere related genes mutations and *MUC5B* rs35705950 T risk allele for all patients, had major contribution in the interpretation of data for all patients and wrote parts of the manuscript; RB had major contribution in the interpretation of data for all patients and revised critically this work for important intellectual content; IB contributed significantly to the genetic analysis regarding the telomere related genes mutations and *MUC5B* rs35705950 T risk allele for all patients and to the interpretation of data; DB, AM, KA, LK had major contribution to the acquisition and interpretation of data for the work and revised critically this work for important intellectual content; AIP performed the statistical analysis of the data, contributed substantially to the interpretation of data for the work and drafted parts of the manuscript and the figures; AT, ZD, EF, DP, AX, AK, SL, IK, IPT, AKK, AG, PS, ICP, CB, AH, TV, MM, KK, EM, EB, CR, GK, AT, ID, TK, PL, NM, SS had major contribution to the acquisition and analysis of data, PR and EL had major contribution to the measurement of telomere length for patients and to the critical interpretation of data; PD revised critically this work for important intellectual content; KR, CM, CO, CM and AP contributed substantially to the analysis and interpretation of data for the work; ML performed the genetic analysis regarding the surfactant protein related genes mutations, had major contribution in the interpretation of data for all patients and wrote parts of the manuscript; NN, ACLH, AC, SA, CB had major contribution in the interpretation of data for all patients and revised critically this work for important intellectual content;

BC had major contribution to the concept and design of the study, to the analysis and interpretation of data and revised critically this work for important intellectual content; SAP and EDM had major contribution to the concept and design of the study, to the acquisition, analysis and interpretation of data, had access to all data, supervised the accuracy and integrity of any part of the work and wrote the manuscript. All authors read and approved of the final version of the submitted publication. EDM and CK had contributed equally to this work. BC and SAP had contributed equally to this work.

Data Availability Statement

Data on which the conclusions of the paper rely are available to editors, reviewers and readers without unnecessary restriction wherever possible.

Figure legends

Figure 1. Pie chart displaying the proportions of different genotype groups in the national Greek IPF cohort with suspected inheritable pulmonary fibrosis.

Figure 2. Kaplan-Meier curves to the time-to-death among different genotype groups and more precisely between patients with telomere related gene (TRG) pathogenic variations alone, patients with TRG pathogenic variations and *MUC5B* rs35705950 T risk allele, patients with neither TRG pathogenic variations nor *MUC5B* rs35705950 T risk allele and patients with *MUC5B* rs35705950 T risk allele alone ($p=0.025$). No death was reported in the *ABCA3* pathogenic variation group and therefore these patients are not included in the analysis. Censoring is indicated by the cross marking. The p-value is derived by log-rank test.

References

- 1 Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, 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 Sep 1;198(5): e44-e68. 8
- 2 Deutsch GH, Young LR, Deterding RR, Fan LL, Dell SD, Bean JA, et al; ChILD Research Co-operative. Diffuse lung disease in young children: application of a novel classification scheme. *Am J Respir Crit Care Med*. 2007 Dec 1;176(11):1120-8.
- 3 Steele MP, Speer MC, Loyd JE, Brown KK, Herron A, Slifer SH, et al. Clinical and pathologic features of familial interstitial pneumonia. *Am J Respir Crit Care Med* 2005 Nov 1;172(9):1146-52
- 4 Alder JK, Hanumanthu VS, Strong MA, DeZern AE, Stanley SE, Takemoto CM, et al. Diagnostic utility of telomere length testing in a hospital-based setting. *Proc Natl Acad Sci U S A*. 2018 Mar 6; 115(10): E2358-E2365.
- 5 Noguee LM, Dunbar AE 3rd, Wert SE, Askin F, Hamvas A, Whitsett JA. A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N Engl J Med*. 2001 Feb 22;344(8):573-9
- 6 Armanios MY, Chen JJ, Cogan JD, Alder JK, Ingersoll RG, Markin C, et al. Telomerase mutations in families with idiopathic pulmonary fibrosis *N Engl J Med*. 2007 Mar 29;356(13):1317-26

7 Tsakiri KD, Cronkhite JT, Kuan PJ, Xing C, Raghu G, Weissler JC, et al. Adult-onset pulmonary fibrosis caused by mutations in telomerase Proc Natl Acad Sci U S A. 2007 May 1;104(18):7552-7

8 Seibold MA, Wise AL, Speer MC, Steele MP, Brown KK, Loyd JE, et al. A common MUC5B promoter polymorphism and pulmonary fibrosis N Engl J Med. 2011 Apr 21;364(16):1503-12

9 Kropski JA, Lawson WE, Young LR, Blackwell TS. Genetic studies provide clues on the pathogenesis of idiopathic pulmonary fibrosis. Dis Model Mech. 2013 Jan;6(1):9-17.

10 Schwartz DA. Idiopathic Pulmonary Fibrosis Is a Genetic Disease Involving Mucus and the Peripheral Airways. Ann Am Thorac Soc. 2018 Nov;15(Suppl 3):S192-S197

11 Nathan N, Giraud V, Picard C, Nunes H, Dastot-Le Moal F, Copin B, et al. Germline SFTPA1 mutation in familial idiopathic interstitial pneumonia and lung cancer. Hum Mol Genet. 2016 Apr 15;25(8):1457-67.

12 Legendre M, Butt A, Borie R, Debray MP, Bouvry D, Filhol-Blin E, et al. Functional assessment and phenotypic heterogeneity of *SFTPA1* and *SFTPA2* mutations in interstitial lung diseases and lung cancer. Eur Respir J. 2020 Dec 24;56(6):2002806.

13 Bullard JE, Wert SE, Whitsett JA, Dean M, Nogee LM. ABCA3 mutations associated with pediatric interstitial lung disease. Am J Respir Crit Care Med. 2005 Oct 15;172(8):1026-31.

-
- 14 Borie R, Bouvry D, Cottin V, Gauvain C, Cazes A, Debray MP, et al. Regulator of telomere length 1 (RTEL1) mutations are associated with heterogeneous pulmonary and extra-pulmonary phenotypes. *Eur Respir J*. 2019 Feb 7;53(2):1800508.
- 15 Benyelles M, O'Donohue MF, Kermasson L, Lainey E, Borie R, Lagresle-Peyrou C, et al. NHP2 deficiency impairs rRNA biogenesis and causes pulmonary fibrosis and Høyeraal-Hreidarsson syndrome *Hum Mol Genet*. 2020 Apr 15;29(6):907-922
- 16 Kannengiesser C, Manali ED, Revy P, Callebaut I, Ba I, Borgel A, et al. First heterozygous *NOP10* mutation in familial pulmonary fibrosis. *Eur Respir J*. 2020 Jun 11;55(6):1902465.
- 17 Kropski JA, Reiss S, Markin C, Brown KK, Schwartz DA, Schwarz MI, et al. Rare Genetic Variants in PARN Are Associated with Pulmonary Fibrosis in Families. *Am J Respir Crit Care Med*. 2017 Dec 1;196(11):1481-1484
- 18 Stuart BD, Choi J, Zaidi S, Xing C, Holohan B, Chen R, et al. Exome sequencing links mutations in PARN and RTEL1 with familial pulmonary fibrosis and telomere shortening. *Nat Genet*. 2015 May;47(5):512-7
- 19 Borie R, Tabèze L, Thabut G, Nunes H, Cottin V, Marchand-Adam S, et al. Prevalence and characteristics of TERT and TERC mutations in suspected genetic pulmonary fibrosis *Eur Respir J*. 2016 Dec;48(6):1721-1731.
- 20 Newton CA, Batra K, Torrealba J, Kozlitina J, Glazer CS, Aravena C, et al. Telomere-related lung fibrosis is diagnostically heterogeneous but uniformly progressive *Eur Respir J*. 2016 Dec;48(6):1710-1720.
- 21 Papiris SA, Tsirigotis P, Kannengiesser C, Kolilekas L, Gkirkas K, Papaioannou AI, et al. Myelodysplastic syndromes and idiopathic pulmonary fibrosis: a dangerous liaison *Respir Res*. 2019 Aug 13;20(1):182

-
- 22 Manali ED, Legendre M, Nathan N, Kannengiesser C, Coulomb-L'Hermine A, Tsiligiannis T, et al. Bi-allelic missense ABCA3 mutations in a patient with childhood ILD who reached adulthood. *ERJ Open Res.* 2019 Jul 22;5(3):00066-2019
- 23 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015 May;17(5):405-24
- 24 Borie R, Kannengiesser C, Gouya L, Dupin C, Amselem S, Ba I, et al. Pilot experience of multidisciplinary team discussion dedicated to inherited pulmonary fibrosis. *Orphanet J Rare Dis.* 2019 Dec 3;14(1):280.
- 25 Juge PA, Lee JS, Ebstein E, Furukawa H, Dobrinskikh E, Gazal S, et al. MUC5B Promoter Variant and Rheumatoid Arthritis with Interstitial Lung Disease. *N Engl J Med.* 2018 Dec 6;379(23):2209-2219.
- 26 Chung JH, Peljto AL, Chawla A, Talbert JL, McKean DF, Rho BH, et al. CT Imaging Phenotypes of Pulmonary Fibrosis in the MUC5B Promoter Site Polymorphism *Chest.* 2016 May;149(5):1215-22.
- 27 van der Vis JJ, Snetselaar R, Kazemier KM, ten Klooster L, Grutters JC, van Moorsel CH. Effect of Muc5b promoter polymorphism on disease predisposition and survival in idiopathic interstitial pneumonias. *Respirology.* 2016 May;21(4):712-7.
- 28 Peljto AL, Zhang Y, Fingerlin TE, Ma SF, Garcia JG, Richards TJ, et al. Association between the MUC5B promoter polymorphism and survival in patients with idiopathic pulmonary fibrosis. *JAMA.* 2013 Jun 5;309(21):2232-9.

29 Peljto AL, Selman M, Kim DS, Murphy E, Tucker L, Pardo A, et al. The MUC5B promoter polymorphism is associated with idiopathic pulmonary fibrosis in a Mexican cohort but is rare among Asian ancestries. *Chest*. 2015 Feb;147(2):460-464.

30 Gable DL, Gaysinskaya V, Atik CC, Talbot CC Jr, Kang B, Stanley SE, et al. ZCCHC8, the nuclear exosome targeting component, is mutated in familial pulmonary fibrosis and is required for telomerase RNA maturation. *Genes Dev*. 2019 Oct 1;33(19-20):1381-1396.

31 Dressen A, Abbas AR, Cabanski C, Reeder J, Ramalingam TR, Neighbors M, 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 Aug;6(8):603-614.

32 Flamein F, Riffault L, Muselet-Charlier C, Pernelle J, Feldmann D, Jonard L, et al. Molecular and cellular characteristics of ABCA3 mutations associated with diffuse parenchymal lung diseases in children. *Hum Mol Genet*. 2012 Feb 15;21(4):765-75

33 Wambach JA, Casey AM, Fishman MP, Wegner DJ, Wert SE, Cole FS, et al. Genotype-phenotype correlations for infants and children with ABCA3 deficiency. *Am J Respir Crit Care Med*. 2014 Jun 15;189(12):1538-43

34 Kröner C, Wittmann T, Reu S, Teusch V, Klemme M, Rauch D, et al. Lung disease caused by ABCA3 mutations. *Thorax*. 2017 Mar;72(3):213-220

35 Liptzin DR, Watson AM, Murphy E, Kroehl ME, Dishop MK, Galambos C, et al. MUC5B expression and location in surfactant protein C mutations in children. *Pediatr Pulmonol*. 2015 Dec;50(12):1270-6.

36 Diaz de Leon A, Cronkhite JT, Katzenstein AL, Godwin JD, Raghu G, Glazer CS, et al. Telomere lengths, pulmonary fibrosis and telomerase (TERT) mutations PLoS One. 2010 May 19;5(5):e10680.

37 Scott MKD, Quinn K, Li Q, Carroll R, Warsinske H, Vallania F, et al. Increased monocyte count as a cellular biomarker for poor outcomes in fibrotic diseases: a retrospective, multicentre cohort study. Lancet Respir Med. 2019 Jun;7(6):497-508.

38 Dudbridge F, Allen RJ, Sheehan NA, Schmidt AF, Lee JC, Jenkins RG, et al. Adjustment for index event bias in genome-wide association studies of subsequent events. Nat Commun. 2019 Apr 5 ;10(1) :1561.

39 Borie R, Kannengiesser C, Hirschi S, Le Pavec J, Mal H, Bergot E, et al ; Groupe d'Etudes et de Recherche sur les Maladies "Orphelines" Pulmonaires (GERM"O" P). Severe hematologic complications after lung transplantation in patients with telomerase complex mutations. J Heart Lung Transplant. 2015 Apr;34(4):538-46.

40 Spagnolo P, Cottin V. Genetics of idiopathic pulmonary fibrosis: from mechanistic pathways to personalized medicine. Med Genet. 2017 Feb;54(2):93-99

