



HAL
open science

Functional assessment and phenotypic heterogeneity of SFTPA1 and SFTPA2 mutations in interstitial lung diseases and lung cancer

Marie Legendre, Afifaa Butt, Raphaël Borie, Marie-Pierre Debray, Diane Bouvry, Emilie Filhol-Blin, Tifenn Desroziers, Valérie Nau, Bruno Copin, Florence Dastot-Le Moal, et al.

► To cite this version:

Marie Legendre, Afifaa Butt, Raphaël Borie, Marie-Pierre Debray, Diane Bouvry, et al.. Functional assessment and phenotypic heterogeneity of SFTPA1 and SFTPA2 mutations in interstitial lung diseases and lung cancer. *European Respiratory Journal*, 2020, 56 (6), pp.2002806. 10.1183/13993003.02806-2020 . inserm-03794264v2

HAL Id: inserm-03794264

<https://inserm.hal.science/inserm-03794264v2>

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.

1 **Functional assessment and phenotypic heterogeneity of *SFTPA1* and *SFTPA2* mutations**
2 **in interstitial lung diseases and lung cancer**

3

4 Marie Legendre^{1,2*}, Afifaa Butt^{*1}, Raphaël Borie³, Marie-Pierre Debray⁴, Diane Bouvry⁵,
5 Emilie Filhol-Blin², Tifenn Desroziers¹, Valérie Nau², Bruno Copin², Florence Dastot-Le
6 Moal², Mélanie Héry¹, Philippe Duquesnoy¹, Nathalie Allou⁶, Anne Bergeron⁷, Julien
7 Bermudez⁸, Aurélie Cazes⁹, Anne-Laure Chene¹⁰, Vincent Cottin¹¹, Bruno Crestani⁴, Jean-
8 Charles Dalphin¹², Christine Dombret⁴, Bérénice Doray¹³, Clairelyne Dupin⁷, Violaine
9 Giraud¹⁴, Anne Gondouin¹², Laurent Gouya⁷, Dominique Israël-Biet¹⁵, Caroline
10 Kannengiesser¹⁶, Aurélie Le Borgne¹⁷, Sylvie Leroy¹⁸, Elisabeth Longchamp¹⁹, Gwenaël
11 Lorillon⁷, Hilario Nunes⁵, Clément Picard²⁰, Martine Reynaud-Gaubert⁸, Julie Traclet¹¹, Paul
12 de Vuyst²¹, Aurore Coulomb L’Hermine²², Annick Clement^{1,23}, Serge Amselem^{1,2**}, Nadia
13 Nathan^{1,23**}

14 * and **: both authors contributed equally

15

16

17

18

19

20

21

22

23

24

25

1. Sorbonne University, Inserm, Childhood genetic disorders, Paris, France
2. Departement of Genetics, Armand Trousseau Hospital, Assistance Publique Hôpitaux de Paris (APHP), Paris, France
3. Pulmonology Department A, Bichat Hospital; APHP, Université de Paris; Paris, France
4. Radiology Department, Bichat Hospital; APHP, Université de Paris; Paris, France
5. Pulmonology Department, EA 2363, Avicenne Hospital, APHP; Paris 13 University, COMUE Sorbonne Paris Cité, Bobigny, France
6. Pulmonology Department, Felix Guyon Hospital ; Saint Denis de La Reunion, France
7. Pulmonology Department, Saint Louis Hospital ; Université de Paris; Paris, France

- 26 8. Pulmonology Department and Lung Transplant Team ; North Hospital – Assistance
27 Publique Hôpitaux de Marseille (APHM), Marseille– MEPHI, IHU
28 Méditerranée Infection, Aix-Marseille University, France
- 29 9. Pathology Department, Bichat Hospital; APHP, Université de Paris; Paris, France
- 30 10. Pulmonology Department, University Hospital; Nantes, France
- 31 11. Pulmonology Department and Coordinating Reference center for rare pulmonary
32 diseases OrphaLung, Hospices Civils de Lyon, Claude Bernard University Lyon 1;
33 Lyon, France
- 34 12. Pulmonology Department, UMR-CNRS Chrono-Environnement 6249 ; CNRS and
35 CHU Besançon, France
- 36 13. Genetic Department, Felix Guyon Hospital ; Saint Denis de La Reunion, France
- 37 14. Pulmonology Department, Ambroise Paré Hospital, APHP, Boulogne Billancourt,
38 France
- 39 15. Pulmonology Department, Georges Pompidou European Hospital, APHP ; Université
40 de Paris; Paris, France
- 41 16. Genetic Departement, Bichat Hospital; APHP, Université de Paris; Paris, France
- 42 17. Pulmonology Department, Larrey Hospital, Toulouse, France
- 43 18. Pulmonology Department, Pasteur Hospital, Nice, France
- 44 19. Pathology Department, Foch Hospital, Suresnes, France
- 45 20. Pulmonology Department, Foch Hospital, Suresnes, France
- 46 21. Pulmonology Department, Erasme Hospital, Brussels, Belgium
- 47 22. Pathology Department, Armand Trousseau Hospital ; Paris, France
- 48 23. Pediatric Pulmonology Department and Reference center for rare lung diseases
49 RespiRare; Armand Trousseau Hospital; Paris, France

50

51 **Running Title:** Interstitial lung diseases due to *SFTPA1* and *SFTPA2* mutations

52

53 **Take home message 255/256-character (including spaces)**

54 *SFTPA1* and *SFTPA2* mutations lead to similar alterations in SP-A secretion and lung tissue
55 expression. They are associated with a highly variable phenotypic expression ranging from
56 incomplete penetrance to severe interstitial lung diseases and lung cancer.

57

58 **Corresponding author**

59 Dr Nadia Nathan

60 Pediatric Pulmonology Department

61 Sorbonne Université and Hôpital Trousseau AP-HP

62 26 avenue du Dr Arnold Netter, 75012-Paris, France

63 Tel : +33(1) 44 73 66 18

64 Fax : +33(1) 44 73 67 18

65 nadia.nathan@aphp.fr

66

67 **Abstract word count:** 248/250

68 **Text word count:** 3813/3000

69 **Tables and Figures:** 8/8

70 **Supplementary table:** 2

71 **Supplementary figures:** 5

72

73 **ABBREVIATIONS**

74	A549	Adenocarcinomic human alveolar basal epithelial
75	ACMG	American College of Medical Genetics
76	AEC	Alveolar epithelial cell
77	ALAT	Latin American thoracic society
78	ATS	American thoracic society
79	CRD	Carbohydrate recognition domain
80	ER	Endoplasmic reticulum
81	ERS	European respiratory society
82	EV	Empty vector
83	GGO	Ground glass opacities
84	GRCh37	Genome Reference Consortium Human Build 37
85	HEK293T	Human embryonic kidney 293 with derived large T antigen
86	HRCT	High resolution computed tomography
87	HSP	Hypersensitivity pneumonitis
88	ILD	Interstitial lung disease
89	IPF	Idiopathic pulmonary fibrosis
90	JRS	Japanese respiratory society
91	MDT	Multidisciplinary team
92	NSIP	Non-specific interstitial pneumonia
93	NT	Non transfected
94	PF	Pulmonary fibrosis
95	PFT	Pulmonary function tests
96	PPFE	Pleuroparenchymal fibroelastosis
97	RespiFIL	National network for rare lung diseases (www.respiFIL.fr)
98	Rs	Referenced single-nucleotide polymorphism
99	SP	Surfactant protein
100	TGF	Transforming growth factor
101	UIP	Usual interstitial pneumonia
102	UPR	Unfolded protein response
103	WT	Wild type
104		

105 **ABSTRACT**

106 **Introduction:** Interstitial lung diseases (ILD) can be caused by mutations in the *SFTPA1* and
107 *SFTPA2* genes, which encode the surfactant protein complex (SP)-A. Only 11 *SFTPA1/2*
108 mutations have so far been reported worldwide, of which 5 have been functionally assessed. In
109 the framework of ILD molecular diagnosis, we identified 14 independent patients with
110 pathogenic *SFTPA1* or *SFTPA2* mutations. The present study aimed to functionally assess the
111 11 different mutations identified and to accurately describe the disease phenotype of the patients
112 and their affected relatives.

113 **Methods:** The consequences of the 11 *SFTPA1/2* mutations were analyzed both *in vitro* by
114 studying the production and secretion of the corresponding mutated proteins and *ex vivo* by
115 analyzing SP-A expression on lung tissue samples. The associated disease phenotypes were
116 documented.

117 **Results:** For the 11 identified mutations, protein production was preserved, but secretion was
118 abolished. The expression pattern of lung SP-A, available in 6 patients, was altered. The family
119 history reported ILD and/or lung adenocarcinoma in 13/14 (93%) families. Among the 28
120 *SFTPA1/2* mutation carriers, the mean age at ILD onset was 45 [0.6-65] years and 48% of them
121 underwent lung transplantation (mean age 51); 7 carriers were asymptomatic.

122 **Discussion:** This study, which expands the molecular and clinical spectrum of SP-A disorders,
123 shows that those pathogenic *SFTPA1/A2* mutations share similar consequences on SP-A
124 secretion in cell models and lung tissue immunostaining, whereas they are associated with a
125 highly variable phenotypic expression of the disease, ranging from severe forms requiring lung
126 transplantation to incomplete penetrance.

127

128

129

130 **KEY WORDS**

131 Surfactant protein A, interstitial lung disease, idiopathic interstitial pneumonia, pulmonary

132 fibrosis, lung cancer, genetics.

133 INTRODUCTION

134

135 Chronic interstitial lung diseases (ILD) include a heterogeneous group of diffuse parenchymal
136 lung disorders that can affect patients of all ages [1–3]. In adult patients, the most severe form
137 of ILD, but also the most frequent, is idiopathic pulmonary fibrosis (IPF). The clinical course
138 of ILD patients is highly variable and unpredictable [4–6]. An unknown part of this
139 heterogeneity implicates genetic factors [7–9]. Pathogenic mutations have been identified in
140 approximately 30% of familial ILD, mainly in genes encoding the telomerase complex in adult
141 patients, and mainly in genes encoding proteins of the surfactant metabolism (*SFTPA1*,
142 *SFTPA2*, *SFTPB*, *SFTPC*, *ABCA3*, *NKX2-1*) in pediatric cases [7, 10, 8, 11–13].

143 Information on *SFTPA1* and *SFTPA2* defects in ILD remains limited. These genes are closely
144 related paralogs containing 6 exons each. They encode the surfactant proteins (SP)-A1 and SP-
145 A2, two collectins that oligomerize in a “bouquet-flower” octadecamer to generate the SP-A
146 complex. SP-A plays a structural role in the surfactant tubular myelin formation, but also an
147 immunomodulatory role through its carbohydrate recognition domain (CRD) that is encoded
148 by the last exon of *SFTPA1/A2* (i.e. exon 6). The molecular epidemiology of *SFTPA1/A2* is still
149 largely unknown. Indeed, over the last few years, only 11 *SFTPA1/A2* mutations have been
150 reported in patients with familial forms of ILD and lung cancer. The pathogenicity of these
151 mutations was attested by functional tests for only 5 of them [11, 14–20].

152 In the present study, performed in the framework of ILD molecular diagnosis, we identified 11
153 missense *SFTPA1* (n=3) or *SFTPA2* (n=8) class 4 (likely pathogenic) and 5 (pathogenic)
154 mutations - according to the American College of Medical Genetics (ACMG) classification of
155 sequence variations - in 14 independent probands [21]. These findings prompted us to assess
156 the functional consequences of these missense variations by means of *in vitro* studies and to
157 accurately describe the associated disease phenotypes in the 14 involved families.

158

159 **PATIENTS AND METHODS**

160 **Patients**

161 Patients with an ILD and with no telomerase complex mutations, were referred to us in the
162 frame of molecular diagnosis. ILD diagnosis was assessed by expert clinicians of one of the 23
163 adult centres of the French reference network for rare lung diseases (RespiFIL) that covers the
164 whole of France. Phenotypic information - including age at ILD onset, associated extra-
165 respiratory features, lung cancer association, and family history - was collected. Available high-
166 resolution computed tomography (CT)-scans and histologic samples were centrally reviewed
167 [3].

168 When a class 4 or class 5 mutation was identified, a genetic counseling was performed and a
169 familial genetic screening proposed to the symptomatic individuals and the asymptomatic major
170 relatives [21]. Each family was discussed within the French multidisciplinary team meeting for
171 genetic ILD [22]. The study was approved by the relevant ethics committees (“Comité de
172 Protection des Personnes”) and written informed consent was obtained from all participants or
173 legal representatives. Clinical information was collected in a legally authorized database (CNIL
174 N°681248).

175

176 **Molecular analyses**

177 *SFTPA1* and *SFTPA2* molecular analyses were performed on DNA extracted from peripheral
178 blood leukocytes. The NM_005411.5 and NM_001098668.2 isoforms were used as a reference
179 for *SFTPA1* and *SFTPA2*, respectively. Coding exons and flanking intronic junctions of
180 *SFTPA1* and *SFTPA2* were sequenced with a custom targeted-capture panel (SeqCap EZ
181 Choice, Roche Diagnostics). The library was prepared following manufacturer’s instructions.
182 Data were analyzed through an in-house double pipeline based on Bowtie2 and BWA tools.

183 Reads were visualized with the IGV viewer (Broad Institute). Copy number variation analysis
184 was performed with a depth-ratio comparison between subjects sequenced in the same run.
185 Class 4 (n=10) and class 5 (n=1) mutations were checked by Sanger sequencing using the
186 previously described protocol [19].

187

188 **Haplotype analyses**

189 Haplotype analyses for families 5, 6, and 11 to 14 were performed by combining *SFTPA2* and
190 flanking *SFTPA1* SNP genotyping by targeted-capture in probands and Sanger sequencing in
191 probands and relatives (*SFTPA1* is the closest gene to *SFTPA2*, 50.5kb downstream on
192 chromosome 10). Phasing was performed through allele segregation within each family and
193 through proband NGS data when contiguous variations were included in at least four common
194 sequencing reads.

195

196 **Structural and electrostatic changes induced by the SP-A1 and SP-A2 mutations on the** 197 **CRD**

198 3D structures of wild type (WT) and mutant CRD of SP-A1/A2 were generated using the rat
199 SP-A crystal structure (residues 201-248) from the Protein Data Bank (PDB ID: 3PAR) as a
200 template. Briefly, homology modeling was performed using Modeller software v9.10.
201 Accuracy of output structures was further assessed using Procheck v3.5.4 [23, 24].
202 Conformational changes were evaluated by a RMSD value comparison between the WT and
203 the mutated SP-A1/A2 domains using SuperPose (Back Bone) v1.0 [25]. A <2 RMSD value is
204 in favor of a very similar protein structure. Electrostatic modeling of the protein domain surface
205 was generated with the use of the PyMOL vacuum electrostatics (v0.99).

206

207 **Plasmid constructs**

208 The WT *SFTPA1* and WT *SFTPA2* cDNAs, obtained from human universal RNA, were
209 inserted into the pcDNA3.1_V5_His_TOPO expression vector (Invitrogen, Carlsbad, CA). A
210 27-aminoacid HA-tag and a 24-aminoacid FLAG-tag were subsequently added after the signal
211 peptide to generate the pSFTPA1_WT and pSFTPA2_WT plasmids respectively. Site-directed
212 mutagenesis was performed to generate the mutant plasmids carrying the identified mutations
213 using the Quick Change Site Mutation Mutagenesis kit (New England Biolabs, Ipswich, MA,
214 E0554S). The resulting plasmid constructs were checked by Sanger sequencing of the inserts
215 and cloning site flanking junctions.

216

217 **Cell culture and transfection**

218 WT and mutant pSFTPA1 and pSFTPA2 expression plasmids were used to assess protein
219 production (cell lysate) and secretion (medium) after transfection in HEK293T cells. 600,000
220 HEK293T cells were first cultured at 37°C with 5% CO₂ into 35mm wells in complete medium
221 (DMEM (1X) + GlutaMAX™-I (4.5g/L D-Glucose) (Life Technologies, Paisley, UK 31966-
222 021), 10% fetal bovine serum (FBS), 1% Penicillin/Streptomycin). On day 2, at 80%
223 confluence, the cells were transfected with 1µg of either pSFTPA1_WT, pSFTPA2_WT, the
224 corresponding mutated plasmids or the empty vector by the FuGENE HD (Promega, Madison,
225 WI E2312) method with a 4/1 ratio. At day 4, the cells were refed with complete medium
226 without FBS. Protein production was studied at day 5.

227

228 **Western blot analyses**

229 Aliquots of the cellular lysates and of the cell culture medium were analyzed by western
230 blotting. The membranes were incubated overnight at 4°C, either with a monoclonal anti-HA
231 peroxidase high affinity antibody (Roche 12013819001®, 1/500) or with a monoclonal anti-
232 FLAG peroxidase high affinity antibody (Sigma-Aldrich A8592®, 1/500). Tubulin was used as

233 a loading control of protein expression by using an anti- α -tubulin peroxidase-coupled antibody
234 (Cell Signaling (11H10) 9099®, 1/1,000) incubated overnight. LAMC1 secretion was used as
235 a loading control of protein secretion by using an anti-LAMC1 antibody (Sigma-Aldrich
236 HPA001908®, rabbit, 1/1,000) incubated overnight, followed by a 1-hour incubation with an
237 anti-rabbit peroxidase-coupled antibody (Cell Signaling 7074®, 1/5000). Results are
238 representative of three independent experiments.

239

240 **Immunohistochemistry on lung tissues**

241 Tissue samples from lung biopsies or lung explants were centrally analyzed by optic
242 microscopy coupled with computerized image analysis of stained materials, and studied by
243 haematoxylin and eosin (HE) staining and immunohistochemistry [26]. A mouse SP-A
244 monoclonal antibody (Abcam-32E12, 1/200) was used according to manufacturer's protocol.
245 Histopathological descriptions were provided according to the ATS/ERS/JRS/ALAT Clinical
246 Practice Guideline and SP-A staining was analyzed on the available tissues [3].

247

248

249 **RESULTS**

250

251 **Spectrum of *SFTPA1* and *SFTPA2* mutations in 14 independent patients**

252 A total of 11 heterozygous class 4 or class 5 mutations were identified: 3 *SFTPA1* mutations in
253 three probands and 8 *SFTPA2* mutations in 11 probands (**Table 1**). Eight of them have so far
254 never been reported. All are missense mutations identified in exon 6 that encodes the CRD of
255 SP-A1 and SP-A2. All these mutations involve an invariant or highly conserved residue in SP-
256 A1 and SP-A2 and throughout evolution (**Figure 1**). None of them are prone to alter splicing

257 according to the MaxEntScan tool. Their pathogenicity, as assessed through *in silico* means, is
258 provided in **Table 1** [21]. Three *SFTPA2* mutations have been each characterized in two distinct
259 families: V178M in families 5 and 6, C238S in families 11 and 12, and R242Q in families 13
260 and 14 (**Figure 2**). This could be related to a common ancestor to two families, or to recurrent
261 mutational events. Families 5 and 6 originate respectively from China and North Africa, and
262 families 13 and 15 from North Africa and La Réunion Island (**Table 2**). Haplotyping showed
263 that V178M arose on different haplotypes in families 5 (haplotype H2) and 6 (haplotype H4)
264 (**Supplemental Figure 1A**). The R242Q also arose on distinct haplotypes in families 13
265 (haplotype H11) and 14 (haplotype H14) (**Supplemental Figure 1C**). These data, together with
266 the diverse geographic origins of the families and the fact that those two mutations are G to A
267 transitions involving a CG dimer, strongly support the hypothesis of recurrent mutational events
268 in discrete populations. Regarding the C238S mutation, haplotype analysis of 17 SNPs did not
269 exclude the existence of an ancestor common to families 11 and 12 that both originate from
270 North Africa (Algeria) (**Supplemental Figure 1B**). This variation is not described in the
271 gnomAD database and is a G to C transversion (mutation mechanism independent of CG dimer
272 hotspots). However, the H6 haplotype characterized in both families is made of the most
273 frequent versions of the 17 SNPs in humans, and the occurrence of a recurrent mutational event
274 in a common haplotype cannot be fully rejected.

275

276 **Structural and electrostatic changes induced by the identified SP-A1 and SP-A2** 277 **mutations**

278 As compared to WT SP-A1 and SP-A2 CRD domains, prediction of the 3D structure of the
279 protein carrying the mutations disclosed no major consequences on the overall domain
280 structure: RMSD values < 1.1 between WT and mutated SP-A1 or SP-A2 domains. The

281 electrostatic modeling of the mutant protein domain surface shows polarity changes for 6 of the
282 mutants, as compared to the corresponding WT SP-A CRD domain (**Supplemental Figure 2**).

283

284 **Impact of the identified missense variants on SP-A1/SP-A2 production and secretion**

285 As shown on the western blots performed on whole cell lysates, and on the histograms with
286 densitometry ratio from 3 or more independent experiments (**Figure 3**), the expression of each
287 mutant SP-A1 and SP-A2 protein (35-37 KDa) is similar to the normal proteins in HEK293T
288 lysates ($p > 0.05$, Student test), thereby indicating that the identified *SFTPA1* and *SFTPA2*
289 missense variations have no effect on protein expression. However, the expression of each
290 mutant SP-A1 and SP-A2 protein is dramatically reduced in the cell medium ($p < 0.05$, Student
291 test), thereby demonstrating that these missense variations result in the absence of SP-A1 and
292 SP-A2 secretion and are therefore deleterious.

293

294 **Table 1: Description of the 11 identified *SFTPA1* and *SFTPA2* mutations**

Gene	Mutation name (cDNA)	Mutation name (protein)	Mutation abbreviated name	rs	GnomAD* allele frequency	Conservation**	Prone to alter splicing †	Polyphen score HumDiv ★	Sift score ‡	Mutation class ◆	First report
<i>SFTPA1</i>	c.532G>A	p.(Val178Met)	V178M	1215316727	1/250994	5/7	No	1	0	4	<i>Doubková et al. 2019</i>
	c.631T>C	p.(Trp211Arg)	W211R	_	Not described	7/7	No	1	0	5	<i>Nathan et al. 2016</i>
	c.673G>A	p.(Val225Met)	V225M	_	Not described	7/7	No	1	0	4	This study
<i>SFTPA2</i>	c.512A>T	p.(Asn171Ile)	N171I	_	Not described	7/7	No	1	0	4	This study
	c.532G>A	p.(Val178Met)	V178M	371035540	3/250974	5/7	No	1	0	4	<i>Coghlan et al. 2014</i>
	c.542A>G	p.(Tyr181Cys)	Y181C	-	Not described	7/7	No	0.99	0.01	4	This study
	c.697T>A	p.(Trp233Arg)	W233R	-	Not described	7/7	No	1	0	4	This study
	c.698G>T	p.(Trp233Leu)	W233L	-	Not described	7/7	No	1	0	4	This study
	c.699G>C	p.(Trp233Cys)	W233C	-	Not described	7/7	No	1	0	4	This study
	c.713G>C	p.(Cys238Ser)	C238S	-	Not described	7/7	No	1	0	4	This study
	c.725G>A	p.(Arg242Gln)	R242Q	759988686	3/251444	7/7	No	1	0.05	4	This study

295

296 All identified mutations are located in exon 6 of *SFTPA1* or *SFTPA2*, encoding the carbohydrate domain of the protein.

297 * GnomAD: includes non-filtered and filtered variants (number of mutated alleles / total number of alleles)

298 ** The conservation has been evaluated through alignment of protein sequences of SP-A (SP-A1 and SP-A2) from species (*Homo sapiens*, *Pan troglodytes*,
299 *Canis lupus*, *Bos taurus*, *Mus musculus*, *Rattus norvegicus*, *Gallus gallus*)

300 † According to the MaxEntScan tool that is based on the maximum entropy principle applied to RNA splicing signals.

301 ★ Polyphen HumDiv score is compiled from all damaging alleles with known effects on the molecular function causing human Mendelian diseases, present in
302 the UniProtKB database, together with differences between human proteins and their closely related mammalian homologs, assumed to be non-damaging. A
303 score ≥ 0.957 is considered as probably damaging, a score between 0.453 and 0.956 is considered as possibly damaging and a score ≤ 0.452 is considered as
304 benign.

305 ‡ SIFT predicts whether an amino acid substitution affects protein function based on sequence homology and the physical properties of amino acids.

306 ◆ According to the international guidelines: pathogenic (class 5), likely pathogenic (class 4), prior to this study's functional assessment [21].

307 **Abbreviations:** GRCh37, Genome Reference Consortium Human Build 37; rs, referenced single-nucleotide polymorphism; Polyphen, Polymorphism
 308 Phenotyping v2.

309 **Table 2: Clinical characteristics of the probands with identified *SFTPA1* or *SFTPA2* mutations**

Gene	Mutation	Family	Individual *	Gender	Geographic origin	Age at onset (year)	Exposures	Lung cancer	Other clinical features	CT-scan features of ILD	Family history	Treatments	Outcome
<i>SFTPA1</i>	V178M	1 - GM01220	II.2	M	North Africa	57	Tobacco	Yes	Seizures, heart disease	UIP	ILD and lung cancer	Pirfenidone at 64 for 3 months then Nintedanib. No oxygen. No treatment for adenocarcinoma	Death at 68
	W211R	2 - PP005	IV.3	M	Europe	45	Tobacco, silica	Yes	No	Indeterminate pattern	ILD and lung cancer	Anti-fibrosing therapy.	LT at 57 years-old. Alive at 59
	V225M	3 - GM0537	II.1	F	North Africa	49	Pigeons, parakeet, dust	No	No	Indeterminate pattern	ILD	Anti-fibrosing therapy	LT at 54, Alive at 61
			III.1	F		NA	No	No	No	No ILD			Asymptomatic at 32
<i>SFTPA2</i>	N171I	4 - PP375	IV.1	F	Europe	32	No	No	Rheumatoid arthritis	UIP and PPFE	ILD		LT at 37. Alive at 40
	V178M	5 - GM01611	III.1	F	China	33	No	No	Rheumatoid arthritis	Indeterminate pattern and PPFE	ILD		LT at 45. Alive at 48
		6 - GM02482	II.1	F	North Africa	34	Tobacco	No	No	Indeterminate pattern	ILD and lung cancer	No O ₂ , no treatment	Alive at 37
	Y181C	7 - PP354	II.6	M	North Africa	56	Tobacco	No	No	ND	ILD and lung cancer	Anti-fibrosing therapies, cyclophosphamide, steroid pulses.	Death at 63 from acute exacerbation
	W233R	8 - GM03569	IV.4	M	North Africa	34	Tobacco, dust, cannabis	Yes	No	Indeterminate pattern	ILD and lung cancer	Anti-fibrosing therapy (nintedanib) and chemotherapy, O ₂	Death at 41 from lung fibrosis and adenocarcinoma progression
W233L	9 - GM02867	II.1	M	Europe	45	No	No	No	Indeterminate pattern	ILD and lung cancer		LT at 49, alive at 54	

W233C	10 - GM00127	II.1	F	Europe	50	Tobacco, birds	Yes	No	Fibrosing NSIP	No	O ₂ , cyclophosphamide	Death at 64 from lung cancer and sepsis
C238S	11 - GM01161	III.2	M	North Africa	46	Tobacco, cement	Yes	No	ND	ILD and lung cancer	Pirfenidone 47	LT at 49, death at 51
	12 - GM00095	II.1	M	North Africa	26	No	No	Klinefelter syndrome	HSP	Lung cancer	No treatment	Asymptomatic, Alive at 51
R242Q	13 - PP092	II.1	M	North Africa	29	Tobacco	No	No	ND	ND		LT at 37, death at 39 from LT complications.
	14 - GM00077	II.2	M	Reunion Island	51	No	No	No	NSIP	ILD	O ₂ before LT, no anti-fibrosing therapies	LT at 51, death at 51, 6 weeks later from sepsis

310 * According to **Figure 2**

311 **Abbreviations:** ILD, interstitial lung disease, UIP, usual interstitial pneumonia; NSIP, non-specific interstitial pneumonia; PPFE, pleuroparenchymal
312 fibroelastosis; HSP, hypersensitivity pneumonitis; LT, lung transplantation

313

314

315 **Parenchymal expression pattern of SP-A in the lung tissue of probands**

316 The tissue samples from the lung biopsy or explanted lungs of 4 probands (**Figure 4**) showed
317 “indeterminate for usual interstitial pneumonia (UIP)” pattern. The SP-A expression pattern
318 was characterized by a marked SP-A expression in hyperplastic pneumocytes and a
319 discontinuous expression surrounding the alveolar space. Such expression pattern contrasts
320 with the light expression and thin continuous linear layer of SP-A at the alveolar surface of the
321 control specimen [26]. In all patients, a residual SP-A expression was observed in alveolar
322 macrophages, consistently with the fact that the antibody used in these experiments recognizes
323 both SP-A1 and SP-A2 that are 98% identical. The same features were observed on the lung
324 biopsy of 2 affected siblings (**Supplemental Figure 5**).

325

326 **Clinical characteristics of the 14 patients and their 14 relatives carrying a *SFTPA1/2*** 327 **mutation**

328 The detailed phenotypic characteristics of the 14 unrelated probands are provided in **Table 2**.
329 Eight (57%) of them originated from North Africa. Only one proband aged 29 at ILD onset
330 (family 13, II.1) had neither lung cancer nor family history of lung cancer or ILD. Eleven (78%)
331 had a first-degree history of ILD and 9 (64%) had a personal (n=5, 36% assessed by lung biopsy
332 and/or explant analysis) or family (n=7, 50%) history of lung cancer (**Figure 2**). A total of 10
333 proband chest CT-scans could be centrally reviewed by a thoracic radiologist (**Figure 5**) [3].
334 Only 2 patients showed a usual interstitial pneumonia (UIP) pattern [3]. Half of cases were
335 classified indeterminate for UIP and 3 cases suggested either nonspecific interstitial pneumonia
336 or hypersensitivity pneumonitis (HSP). Two patients had apical subpleural consolidations
337 evoking pleuroparenchymal fibroelastosis associated to UIP and indeterminate fibrosis (one
338 case each). In total, predominant ground glass opacities (GGO) were observed in 7 out of 10

339 (70%) patients and lung fibrosis was observed in 7 out of 10 patients (70%). Other inconstant
340 features were basal thickened interlobular septa, micronodules, reticulations, honeycombing
341 and bronchiectasis. The same results were observed on the chest CT-scans of 2 affected siblings
342 **(Supplemental Figure 4)**.

343

344 The family screening led to identify a *SFTPA1* or *SFTPA2* mutation in 14 more individuals
345 **(Figure 2 and Supplemental Table 1)**.

346 Altogether, the present study allowed identifying 28 individuals with a *SFTPA1/A2* mutation.
347 The male to female ratio was 0.65. The median age at onset of the ILD was 45 years old [0.6-
348 65]. Five (22%) patients presented an extra-respiratory manifestation, including two patients
349 with rheumatoid arthritis (9%). Ten patients (43%) declared neither tobacco nor professional
350 exposure. Seven relatives remained asymptomatic until a median age of 39 [19-55] and after a
351 median length of follow-up of 5, range [1-8] years. The penetrance of the disease was
352 incomplete as among them, 2 individuals were aged over 40 (42 and 55 years respectively) and
353 harbored normal pulmonary function tests (PFT) and CT-scan. The other asymptomatic
354 relatives had either abnormal PFT and/or minimal interstitial abnormalities on the CT-scan at
355 ages 18, 24, 32, 33, 39 years-old respectively. The first and last available PFT of the patients
356 are provided in **Supplemental Figure 3**. The evolution of the ILD was severe: 5 patients needed
357 anti-fibrosing therapies and 11 (48%) patients received a lung transplantation at a median age
358 of 51 [37-58] years with a median time from diagnosis of 5 years [0-19]. Regarding lung
359 transplantation, one patient with ILD and surgically treated lung cancer (Family 2, IV.3) benefit
360 from an atypical indication: despite a minimally invasive with a lepidic pattern adenocarcinoma
361 on the contralateral lung almost 10 years after the first one, lung transplantation was decided,
362 with a 2-years favorable outcome to date. For all the patients, the lung transplantation was
363 bilateral, in order to prevent lung fibrosis extension and lung cancer development on the native

364 lung. Nine (32%) patients died at a median age of 51 years old [0.7-68], three of them shortly
365 after lung transplantation (7 days, 6 weeks and 13 months respectively) (**Supplemental Table**
366 **1** and **Figure 6**).

367

368 **DISCUSSION**

369 The present study describes for the first time a large number of functionally assessed pathogenic
370 *SFTPA1* and *SFTPA2* mutations in 14 independent families. Three mutations were identified in
371 *SFTPA1* and 8 in *SFTPA2*, bringing to 5 and 14 respectively the total number of reported
372 *SFTPA1* and *SFTPA2* mutations (**Supplemental Table 2**) [11, 14–19]. Interestingly, all the
373 described mutations are located in exon 6 that encodes the protein CRD, and all but one
374 (p.(Tyr208His)) were found in the heterozygous state.

375 In this study, the pathogenicity of the identified mutations was assessed by *in vitro* functional
376 studies and *ex-vivo* immunostaining showing that the secretion of all the mutant proteins was
377 similarly abolished and that the protein may be sequestered in the alveolar epithelial cells
378 (AEC)-2. Although part of these results rely on a cell model that overexpresses recombinant
379 proteins, such quite unequivocal presentation of the disease at the cellular level contrasts
380 sharply with the highly variable clinical expression of the disease, which ranges from
381 incomplete penetrance to severe forms requiring lung transplantation.

382

383 **Phenotypic heterogeneity**

384 The individuals with *SFTPA1/A2* mutations presented with a heterogeneous clinical phenotype,
385 with no evidence of phenotypic differences between patients with *SFTPA1* mutations and those
386 with *SFTPA2* mutations. The penetrance of the disease was incomplete, with the identification
387 of carriers remaining asymptomatic in adulthood (with normal chest CT-scan and pulmonary
388 function tests). The age of onset was variable, from childhood to elderly [19]. However, in the

389 majority of the patients, the ILD occurred at a younger age (mean 40 years old; SD 14; range
390 [0.6-65]) than the usual age at IPF onset as only one patient experienced a disease onset after
391 60 years-old. The CT-scan analysis gave highly heterogeneous results in terms of ATS/ERS
392 classification. This probably highlights that this classification is maybe not suitable in such
393 cases. As shown in the current study, and as observed in children with surfactant disorders, the
394 GGO pattern represents the most frequent pattern [27]. This observation differs from the typical
395 UIP pattern that is seen in IPF and could be one of the main arguments – on top of the family
396 form of ILD - for suspecting a surfactant disorder in adults. Of interest, no relationship could
397 be established between the severity of the CT images and the mutations. Moreover, various CT
398 patterns were observed within the same family. Altogether, as it is also the case in other causes
399 of genetic ILD, the variability in disease expression was observed among the different
400 individuals, including among those from a same family [22].

401

402 Variability in clinical expression of the disease phenotype is a common feature of most genetic
403 disorders, especially when dominant. In theory, this phenomenon may result from allelic
404 heterogeneity and/or from the influence of environmental or modifying genetic factors [28].
405 Moreover, post-translational modifications of SP-A proteins could at least partly account for
406 the heterogeneous presentation of the clinical phenotype. The role of viruses has been discussed
407 in pediatric patients who develop fatal pulmonary fibrosis [6]. This hypothesis has also been
408 discussed for mutations in other surfactant genes, especially *SFTPC* and *ABCA3* mutations [29,
409 30]. In adults, tobacco smoking and occupational exposures are supposed to be the main triggers
410 in adult cases of IPF [6, 31] but the role of viral infections as a risk factor for pulmonary fibrosis
411 has also been reported by several authors and recently confirmed in a meta-analysis [32]. The
412 interactions between SP-A and various viruses have been studied *in vitro* and in *SFTPA* knock-
413 out mouse models. The viral infections appeared to promote lung fibrosis development, whereas

414 the adjunction of a SP-A treatment allowed a beneficial effect on lung fibrosis progression [33].
415 SP-A binds the pathogens oligosaccharides via its CRD domain, induces their opsonisation
416 and/or phagocytosis and promotes a pro-inflammatory response [34]. Alternatively, the CRD
417 binding to the alveolar macrophages could induce an anti-inflammatory response by enhancing
418 apoptotic cells phagocytosis. This CRD ability to regulate pro- and anti-inflammatory responses
419 is known as the “head or tail hypothesis” depending on the orientation of SP-A binding to
420 alveolar macrophages via a CRD-dependent or CRD-independent mechanism [35, 36]. In this
421 context, molecular alterations of SP-A1/A2 CRD could alter its ability to respond to pathogens
422 exposures and to modulate pulmonary inflammation caused by environmental triggers.

423

424 The observation of a SP-A accumulation in the AEC2 observed on lung tissues sections
425 suggests the production of misfolded proteins in the AEC2 that could induce variable levels of
426 endoplasmic reticulum (ER) stress and promote cell death or cell reprogramming [37]. The
427 unsecreted abnormal SP-A protein could thus be responsible for an unfolded protein response
428 (UPR) challenge. UPR and ER stress have been found increased in case of *SFTPC* mutations
429 [37]. Evidence for increased ER stress as well as abnormal oligomerization and increased
430 proteasome degradation has also been reported for mutant SP-A2 proteins (G231V and F198S
431 mutations) [38]. Similarly, *SFTPA1* mutations have been associated with an increased
432 expression of ER stress and necroptosis markers (Y208H mutation) [39]. Thus, there are
433 compelling evidence for a major role of ER stress in *SFTPA1/A2* mutations, which need to be
434 further confirmed on patients’ lung tissues.

435

436 **Lung cancer association**

437 IPF has been associated with an increased risk of lung cancer, such association being observed
438 in up to 13.5% of the patients with pulmonary fibrosis [40, 41]. In the present study, the

439 association of ILD and lung cancer reaches 36% in the probands, and 64% when including
440 family history, which strongly supports a link between *SFTPA1/A2* CRD mutations and lung
441 cancer development. Interestingly, such association has not been reported for *SFTPC*
442 mutations. Given the severity of the *SFTPC*-associated disease that leads to lung transplantation
443 or to death at a young age, the follow-up may not be long enough to lead to tumorigenesis
444 promotion. ER stress, via UPR, is also linked to tumorigenesis [42]. Increased ER stress can
445 alter tissue homeostasis by interfering with cell cycle in many ways, including cell death or cell
446 senescence attenuation (and thus tumorigenesis promotion) [43]. In case of SP-A mutations,
447 these hypotheses still need to be further evaluated. It has also been proposed that an increased
448 expression of SP-A could reduce adenocarcinogenesis in mice [44]. This could be explained by
449 the interactions between SP-A and Transforming growth factor (TGF)- β , a crucial molecule
450 implicated in lung repair and tumor proliferation: SP-A has been shown to inhibit TGF- β 1
451 inactivation [45, 46]. Thus, mutations of SP-A that alter its secretion could prevent SP-A from
452 inhibiting the TGF- β pathway, leading to an impairment of the alveolar healing process [47].

453

454 **Genetic studies, family screening and genetic counselling**

455 This study suggests that a genetic screening for *SFTPA1/A2* mutations should be proposed to
456 patients with (i) a family form of ILD, especially if associated to a personal or family history
457 of lung cancer; (ii) an early-onset ILD (before 50 years) and no telomerase complex gene
458 mutations. When a mutation is found, in line with the incomplete penetrance and the risk of
459 ILD and lung cancer, a family screening should be offered to the siblings and other relatives,
460 including asymptomatic major relatives. The pre-symptomatic management of individuals with
461 *SFTPA1/A2* mutations is not consensual yet, but it is likely that tobacco smoking and
462 environmental exposures can enhance the disease expression, and that genetic counseling
463 should also include targeted preventive measures. In France, national multidisciplinary team

464 (MDT) meetings dedicated to genetic forms of pulmonary fibrosis have recently been launched
465 in the frame of the RespiFIL network [22]. Among other issues, prenatal and pre-implantation
466 managements are discussed for each family. Following the MDT report, a personalized genetic
467 counseling is given to the patient and his family. As shown in this study, the phenotypic
468 heterogeneity of the disease is important and could change the families' perception of the
469 disease severity and their expectations in terms of prenatal / pre-implantation management.

470

471 **Conclusion**

472 Heterozygous *SFTPA1* and *SFTPA2* mutations involving the exon 6 that encodes the CRD of
473 the corresponding proteins contribute to various fibrotic ILDs that can occur in young adults
474 and probably during childhood, and to a higher risk of lung cancer. Their pathogenicity is
475 attested to by impaired protein secretion leading to cytoplasmic retention in the alveolar
476 epithelium. However, the penetrance of the disease phenotype is incomplete and clinical
477 expression of the disease is highly variable with a major proportion of patients, including young
478 adults, presenting a severe disease requiring lung transplantation. As such, family screening
479 and genetic counseling as well as pre-symptomatic management of the relatives are major issues
480 that need to be discussed in the frame of multidisciplinary team meetings.

481

482 **AUTHORS' CONTRIBUTIONS**

483 NN, ML, AC, SA designed the study; NN, AB, EFB, TD, MH, VN, ACH performed
484 experiments; NN, ML, BC, PD, EFB, TD, MH, AB, ACH, MPD, AC, SA analyzed and
485 interpreted the data; NN and AB wrote the manuscript; ML, SA, AC, BC, MPD, ACH, RB,
486 FDM, CK, LG, VC, JT, BCr, AG, CD, CDo, ACa, ALC, VG, HN, DB, ABe, GL, DIB, PV,
487 CP, EL, MRG, JCD, SL, ALB, NA provided the patients' data; All authors reviewed and
488 approved the manuscript.

489

490 **ACKNOWLEDGMENTS**

491 We wish to thank the patients and their families for their participation in the study. We also
492 thank the Assistance Publique Hôpitaux de Paris, Sorbonne Université, Paris, France, and the
493 national networks for rare lung diseases: Centre de référence des maladies respiratoires rares
494 (RespiRare), Centre de référence des maladies pulmonaires rares (OphaLung) and Filière de
495 soins pour les maladies respiratoires rares (RespiFIL). The ILD cohort is developed in
496 collaboration with the Rare Cohort Disease (RaDiCo)-ILD project (ANR-10-COHO-0003), the
497 FP7-305653-child-EU project and the COST Action European network for translational
498 research in children's and adult interstitial lung disease (COST-ILD) project (CA16125).

499

500 **CONFLICT OF INTEREST STATEMENT**

501 Dr Borie declared he has been invited by to national and international meetings, and/or that he
502 received grants and/or personal fees for various missions from Boehringer Ingelheim, Roche,
503 Savapharma, outside the submitted work.

504 Prof Cottin declared he has been invited by to national and international meetings, and/or that
505 he received grants and/or personal fees for various missions from Actelion, Boehringer
506 Ingelheim, Bayer/MSD, GSK, Novartis, Roche/Promedior, Sanofi, Celgene, Galapagos,
507 outside the submitted work.

508 The other authors declare no conflict of interest.

509

510 **FUNDING**

511 This work was supported by grants from the Institut National de la Santé et la Recherche
512 Médicale (INSERM), the Legs Poix from the Chancellerie des Universités (grants 2013 n°1305,
513 2014 n°1405, 2015 n°1015, 2016 n°2077 and 2017 n°DP2017/1860), Paris, the Société

514 Française de Pédiatrie - Société Pédiatrique de Pneumologie et d'Allergologie- AstraZeneca,
515 as well as fundings from the patient organizations Respirer c'est Grandir and Belleherbe
516 Association.

517

518 REFERENCES

- 519 1. Bush A, Cunningham S, de Blic J, Barbato A, Clement A, Epaud R, Hengst M, Kiper
520 N, Nicholson AG, Wetzke M, Snijders D, Schwerk N, Griese M, chILD-EU collaboration.
521 European protocols for the diagnosis and initial treatment of interstitial lung disease in
522 children. *Thorax* 2015; 70: 1078–1084.
- 523 2. Kurland G, Deterding RR, Hagood JS, Young LR, Brody AS, Castile RG, Dell S, Fan
524 LL, Hamvas A, Hilman BC, Langston C, Nogee LM, Redding GJ, American Thoracic Society
525 Committee on Childhood Interstitial Lung Disease (chILD) and the chILD Research Network.
526 An official American Thoracic Society clinical practice guideline: classification, evaluation,
527 and management of childhood interstitial lung disease in infancy. *Am. J. Respir. Crit. Care*
528 *Med.* 2013; 188: 376–394.
- 529 3. Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, Behr J,
530 Cottin V, Danoff SK, Morell F, Flaherty KR, Wells A, Martinez FJ, Azuma A, Bice TJ,
531 Bouros D, Brown KK, Collard HR, Duggal A, Galvin L, Inoue Y, Jenkins RG, Johkoh T,
532 Kazerooni EA, Kitaichi M, Knight SL, Mansour G, Nicholson AG, Pipavath SNJ, Buendía-
533 Roldán I, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official
534 ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am. J. Respir. Crit. Care Med.* 2018; 198:
535 e44–e68.
- 536 4. Collard HR, Ryerson CJ, Corte TJ, Jenkins G, Kondoh Y, Lederer DJ, Lee JS, Maher
537 TM, Wells AU, Antoniou KM, Behr J, Brown KK, Cottin V, Flaherty KR, Fukuoka J,
538 Hansell DM, Johkoh T, Kaminski N, Kim DS, Kolb M, Lynch DA, Myers JL, Raghu G,
539 Richeldi L, Taniguchi H, Martinez FJ. Acute Exacerbation of Idiopathic Pulmonary Fibrosis.
540 An International Working Group Report. *Am. J. Respir. Crit. Care Med.* 2016; 194: 265–275.
- 541 5. Clement A, de Blic J, Epaud R, Galeron L, Nathan N, Hadchouel A, Barbato A,
542 Snijders D, Kiper N, Cunningham S, Griese M, Bush A, Schwerk N, chILD-EU collaboration.
543 Management of children with interstitial lung diseases: the difficult issue of acute
544 exacerbations. *Eur. Respir. J.* 2016; 48: 1559–1563.
- 545 6. Nathan N, Sileo C, Thouvenin G, Berdah L, Delestrain C, Manali E, Papisiris S, Léger
546 P-L, Pointe HD le, l'Hermine AC, Clement A. Pulmonary Fibrosis in Children. *J. Clin. Med.*
547 2019; 8.
- 548 7. Garcia CK. Idiopathic pulmonary fibrosis: update on genetic discoveries. *Proc Am*
549 *Thorac Soc* 2011; 8: 158–162.
- 550 8. Turcu S, Ashton E, Jenkins L, Gupta A, Mok Q. Genetic testing in children with
551 surfactant dysfunction. *Arch. Dis. Child.* 2013; 98: 490–495.
- 552 9. Nathan N, Borensztajn K, Clement A. Genetic causes and clinical management of
553 pediatric interstitial lung diseases. *Curr. Opin. Pulm. Med.* 2018; 24: 253–259.
- 554 10. Kropski JA, Lawson WE, Young LR, Blackwell TS. Genetic studies provide clues on
555 the pathogenesis of idiopathic pulmonary fibrosis. *Dis. Model. Mech.* 2013; 6: 9–17.
- 556 11. Coghlan MA, Shifren A, Huang HJ, Russell TD, Mitra RD, Zhang Q, Wegner DJ,
557 Cole FS, Hamvas A. Sequencing of idiopathic pulmonary fibrosis-related genes reveals
558 independent single gene associations. *BMJ Open Respir. Res.* 2014; 1: e000057.

- 559 12. Borie R, Kannengiesser C, Sicre de Fontbrune F, Gouya L, Nathan N, Crestani B.
560 Management of suspected monogenic lung fibrosis in a specialised centre. *Eur. Respir. Rev.*
561 2017; 26.
- 562 13. Newton CA, Batra K, Torrealba J, Kozlitina J, Glazer CS, Aravena C, Meyer K,
563 Raghu G, Collard HR, Garcia CK. Telomere-related lung fibrosis is diagnostically
564 heterogeneous but uniformly progressive. *Eur. Respir. J.* 2016; 48: 1710–1720.
- 565 14. Wang Y, Kuan PJ, Xing C, Cronkhite JT, Torres F, Rosenblatt RL, DiMaio JM, Kinch
566 LN, Grishin NV, Garcia CK. Genetic defects in surfactant protein A2 are associated with
567 pulmonary fibrosis and lung cancer. *Am. J. Hum. Genet.* 2009; 84: 52–59.
- 568 15. van Moorsel CHM, Ten Klooster L, van Oosterhout MFM, de Jong PA, Adams H,
569 Wouter van Es H, Ruven HJT, van der Vis JJ, Grutters JC. SFTPA2 Mutations in Familial
570 and Sporadic Idiopathic Interstitial Pneumonia. *Am. J. Respir. Crit. Care Med.* 2015; 192:
571 1249–1252.
- 572 16. Doubková M, Staňo Kozubík K, Radová L, Pešová M, Trizuljak J, Pál K, Svobodová
573 K, Réblová K, Svozilová H, Vrzalová Z, Pospíšilová Š, Doubek M. A novel germline
574 mutation of the SFTPA1 gene in familial interstitial pneumonia. *Hum. Genome Var.* 2019; 6:
575 12.
- 576 17. Takezaki A, Tsukumo S-I, Setoguchi Y, Ledford JG, Goto H, Hosomichi K, Uehara
577 H, Nishioka Y, Yasutomo K. A homozygous SFTPA1 mutation drives necroptosis of type II
578 alveolar epithelial cells in patients with idiopathic pulmonary fibrosis. *J. Exp. Med.* 2019; .
- 579 18. Sritharan SS, Gajewska ME, Skytte A-BS, Madsen LB, Bendstrup E. Familial
580 idiopathic pulmonary fibrosis in a young female. *Respir. Med. Case Rep.* 2018; 24: 1–4.
- 581 19. Nathan N, Giraud V, Picard C, Nunes H, Dastot-Le Moal F, Copin B, Galeron L, de
582 Ligniville A, Kuziner N, Reynaud-Gaubert M, Valeyre D, Couderc L-J, Chinet T, Borie R,
583 Crestani B, Simansour M, Nau V, Tissier S, Duquesnoy P, Mansour-Hendili L, Legendre M,
584 Kannengiesser C, Coulomb L'Hermine A, Gouya L, Amselem S, Clement A. Germline
585 SFTPA1 mutation in familial idiopathic interstitial pneumonia and lung cancer. *Hum. Mol.*
586 *Genet.* 2016; 25: 1457–1467.
- 587 20. Liu L, Qin J, Guo T, Chen P, Ouyang R, Peng H, Luo H. Identification and functional
588 characterization of a novel surfactant protein A2 mutation (p.N207Y) in a Chinese family
589 with idiopathic pulmonary fibrosis. *Mol. Genet. Genomic Med.* 2020; : e1393.
- 590 21. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M,
591 Lyon E, Spector E, Voelkerding K, Rehm HL, ACMG Laboratory Quality Assurance
592 Committee. Standards and guidelines for the interpretation of sequence variants: a joint
593 consensus recommendation of the American College of Medical Genetics and Genomics and
594 the Association for Molecular Pathology. *Genet. Med.* 2015; 17: 405–424.
- 595 22. Borie R, Kannengiesser C, Gouya L, Dupin C, Amselem S, Ba I, Bunel V, Bonniaud
596 P, Bouvry D, Cazes A, Clement A, Debray MP, Dieude P, Epaud R, Fanen P, Lainey E,
597 Legendre M, Plessier A, Sicre de Fontbrune F, Wemeau-Stervinou L, Cottin V, Nathan N,
598 Crestani B. Pilot experience of multidisciplinary team discussion dedicated to inherited
599 pulmonary fibrosis. *Orphanet J. Rare Dis.* 2019; 14: 280.
- 600 23. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to
601 check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 1993; 26: 283–
602 291.
- 603 24. Rullmann JAC. AQUA, Computer Program, Department of NMR Spectroscopy,
604 Bijvoet Center for Biomolecular Research, Utrecht University, The Netherlands. 1996.
- 605 25. Maiti R, Van Domselaar GH, Zhang H, Wishart DS. SuperPose: a simple server for
606 sophisticated structural superposition. *Nucleic Acids Res.* 2004; 32: W590–W594.
- 607 26. Ochs M, Johnen G, Müller K-M, Wahlers T, Hawgood S, Richter J, Brasch F.
608 Intracellular and intraalveolar localization of surfactant protein A (SP-A) in the parenchymal

- 609 region of the human lung. *Am. J. Respir. Cell Mol. Biol.* 2002; 26: 91–98.
- 610 27. Semple TR, Ashworth MT, Owens CM. Interstitial Lung Disease in Children Made
611 Easier...Well, Almost. *Radiogr. Rev. Publ. Radiol. Soc. N. Am. Inc* 2017; 37: 1679–1703.
- 612 28. Nathan N, Corvol H, Amselem S, Clement A. Biomarkers in Interstitial lung diseases.
613 *Paediatr. Respir. Rev.* 2015; 16: 219–224.
- 614 29. Glasser SW, Witt TL, Senft AP, Baatz JE, Folger D, Maxfield MD, Akinbi HT,
615 Newton DA, Prows DR, Korfhagen TR. Surfactant protein C-deficient mice are susceptible to
616 respiratory syncytial virus infection. *Am. J. Physiol. Lung Cell. Mol. Physiol.* 2009; 297: L64-
617 72.
- 618 30. Kaltenborn E, Kern S, Frixel S, Fragnet L, Conzelmann K-K, Zarbock R, Griese M.
619 Respiratory syncytial virus potentiates ABCA3 mutation-induced loss of lung epithelial cell
620 differentiation. *Hum. Mol. Genet.* 2012; 21: 2793–2806.
- 621 31. Vancheri C, Cottin V, Kreuter M, Hilberg O. IPF, comorbidities and management
622 implications. *Sarcoidosis Vasc. Diffuse Lung Dis. Off. J. WASOG World Assoc. Sarcoidosis*
623 *Granulomatous Disord.* 2015; 32 Suppl 1: 17–23.
- 624 32. Sheng G, Chen P, Wei Y, Yue H, Chu J, Zhao J, Wang Y, Zhang W, Zhang H-L. Viral
625 Infection Increases the Risk of Idiopathic Pulmonary Fibrosis: A Meta-Analysis. *Chest* 2020;
626 157: 1175–1187.
- 627 33. Al-Qahtani AA, Murugaiah V, Bashir HA, Pathan AA, Abozaid SM, Makarov E, Nal-
628 Rogier B, Kishore U, Al-Ahdal MN. Full-length human surfactant protein A inhibits influenza
629 A virus infection of A549 lung epithelial cells: A recombinant form containing neck and
630 lectin domains promotes infectivity. *Immunobiology* 2019; 224: 408–418.
- 631 34. Lopez-Rodriguez E, Pascual A, Arroyo R, Floros J, Perez-Gil J. Human Pulmonary
632 Surfactant Protein SP-A1 Provides Maximal Efficiency of Lung Interfacial Films. *Biophys. J.*
633 2016; 111: 524–536.
- 634 35. Atochina-Vasserman EN, Beers MF, Gow AJ. Review: Chemical and structural
635 modifications of pulmonary collectins and their functional consequences. *Innate Immun.*
636 2010; 16: 175–182.
- 637 36. Gardai SJ, Xiao Y-Q, Dickinson M, Nick JA, Voelker DR, Greene KE, Henson PM.
638 By binding SIRPalpha or calreticulin/CD91, lung collectins act as dual function surveillance
639 molecules to suppress or enhance inflammation. *Cell* 2003; 115: 13–23.
- 640 37. Mulugeta S, Nguyen V, Russo SJ, Muniswamy M, Beers MF. A surfactant protein C
641 precursor protein BRICHOS domain mutation causes endoplasmic reticulum stress,
642 proteasome dysfunction, and caspase 3 activation. *Am J Respir Cell Mol Biol* 2005; 32: 521–
643 530.
- 644 38. Maitra M, Wang Y, Gerard RD, Mendelson CR, Garcia CK. Surfactant protein A2
645 mutations associated with pulmonary fibrosis lead to protein instability and endoplasmic
646 reticulum stress. *J Biol Chem* 2010; 285: 22103–22113.
- 647 39. Sano R, Reed JC. ER stress-induced cell death mechanisms. *Biochim. Biophys. Acta*
648 *BBA - Mol. Cell Res.* 2013; 1833: 3460–3470.
- 649 40. JafariNezhad A, YektaKooshali MH. Lung cancer in idiopathic pulmonary fibrosis: A
650 systematic review and meta-analysis. *PLoS One* 2018; 13: e0202360.
- 651 41. Tzouveleakis A, Spagnolo P, Bonella F, Vancheri C, Tzilas V, Crestani B, Kreuter M,
652 Bouros D. Patients with IPF and lung cancer: diagnosis and management. *Lancet Respir. Med.*
653 2018; 6: 86–88.
- 654 42. Zhu B, Ferry CH, Markell LK, Blazanin N, Glick AB, Gonzalez FJ, Peters JM. The
655 nuclear receptor peroxisome proliferator-activated receptor- β/δ (PPAR β/δ) promotes
656 oncogene-induced cellular senescence through repression of endoplasmic reticulum stress. *J.*
657 *Biol. Chem.* 2014; 289: 20102–20119.
- 658 43. Lafont E. Stress Management: Death Receptor Signalling and Cross-Talks with the

659 Unfolded Protein Response in Cancer. *Cancers* 2020; 12.
660 44. Mitsuhashi A, Goto H, Kuramoto T, Tabata S, Yukishige S, Abe S, Hanibuchi M,
661 Kakiuchi S, Saijo A, Aono Y, Uehara H, Yano S, Ledford JG, Sone S, Nishioka Y. Surfactant
662 protein A suppresses lung cancer progression by regulating the polarization of tumor-
663 associated macrophages. *Am. J. Pathol.* 2013; 182: 1843–1853.
664 45. Willems CHMP, Zimmermann LJI, Kloosterboer N, Kramer BW, van Iwaarden JF.
665 Surfactant protein A binds TGF- β 1 with high affinity and stimulates the TGF- β pathway.
666 *Innate Immun.* 2014; 20: 192–199.
667 46. Willems CHMP, Zimmermann LJI, Langen RMR, van den Bosch MJA, Kloosterboer
668 N, Kramer BW, van Iwaarden JF. Surfactant protein A influences reepithelialization in an
669 alveolocapillary model system. *Lung* 2012; 190: 661–669.
670 47. Saito A, Horie M, Micke P, Nagase T. The Role of TGF- β Signaling in Lung Cancer
671 Associated with Idiopathic Pulmonary Fibrosis. *Int. J. Mol. Sci.* 2018; 19.
672

673 FIGURES

674

675 **Figure 1: SP-A1/SP-A2 mutation localization, SP-A1/SP-A2 conservation and interspecies** 676 **conservation of the SP-A1 and SP-A2 amino-acid impacted by the mutations**

677 Reported *SFTPA1* and *SFTPA2* mutations are respectively in green and red (see also **Table 1**
678 and **Supplemental Table 1**). The mutations identified in this study are in bold. The aminoacids
679 of the highly homologous carbohydrate recognition domain of SP-A1, SP-A2 and of 7 species
680 were aligned (*Homo sapiens*, *Pan troglodytes*, *Canis lupus*, *Bos taurus*, *Mus musculus*, *Rattus*
681 *norvegicus*, *Gallus gallus*). * aminoacids invariant between SP-A1 and SP-A2 and throughout
682 evolution.

683

684 **Figure 2: Genealogical trees of the 14 included families with *SFTPA1* or *SFTPA2*** 685 **mutations**

686 The arrows indicate the probands. Individuals with interstitial lung disease (ILD) are indicated
687 by a black symbol. Individuals with lung cancer are indicated by a checkerboard symbol.
688 Symbols that are half black and half checkerboard indicate individuals with ILD and lung
689 cancer. Light grey symbols represent undefined chronic lung diseases reported by the family.
690 Each individual is identified by his generation (Roman numbers) and its number (Arabic

691 number). +/- indicates the presence of a *SFTPA1* or *SFTPA2* mutation in the heterozygous state;
692 -/- indicates an absence of *SFTPA1* or *SFTPA2* mutation.

693

694 **Figure 3: Effects of *SFTPA1* and *SFTPA2* variants on protein production and secretion**
695 **in HEK293T cells**

696 **A.** Western blot. A 35-37-kDa molecular species consistent with the HA-tagged SP-A1 or
697 FLAG-tagged SP-A2 proteins is observed in lysates of cells expressing the wild-type (WT) or
698 mutant SP-A isoforms. A similar 35-37-kDa molecular species is also observed in the medium
699 of cells expressing the WT proteins, but in the medium of the cells transfected with the mutant
700 constructs of SP-A1 or SP-A2, an absence of secretion is observed. α -tubulin production was
701 used as a loading control of protein expression and LAMC1 secretion was used as a loading
702 control of protein secretion.

703 **B.** Normalized protein expression based on densitometry from western blot experiments. Charts
704 show HA-tagged SP-A1 and FLAG-tagged SP-A2 protein expression normalized by α -tubulin
705 (cell lysate expression) and LAMC1 (cell medium expression). Data are means of densitometry
706 ratio (Image J software) from 3 or more independent western blot experiments. Error bars show
707 standard error of the mean.

708 NT: not transfected; WT: wild type; EV: empty vector; *, $p < 0.05$

709

710 **Figure 4: Lung histopathological features and SP-A expression on lung tissue from**
711 **patients heterozygous for a *SFTPA1* or *SFTPA2* mutation**

712 Histological examination was performed on lung tissues from 4 index patients (family 5, III.1;
713 family 3, II.1; family 14, II.1; and family 9, II.1) and compared to an adult lung tissue sample.
714 Haematoxylin and eosin (HE) staining was performed and SP-A expression was analyzed by
715 immunohistochemistry. In all the patients, the HE staining (left column, x25) confirmed the

716 ILD, with however various patterns described according to the 2018 ATS/ERS/JRS/ALAT
717 guideline [3].

718 Tissue from patient F5, III.1 showed diffuse fibrosis without fibroblastic foci delimitating some
719 unaffected lung parenchyma associated with prominent inflammation with lymphoid nodules
720 consistent with indeterminate for UIP.

721 Tissue from patient F3, II.1 showed diffuse fibrosis with very few unaffected lung parenchyma
722 with microscopic honeycombing, inflammation away from areas of honeycombing,
723 bronchiolar metaplasia and osseous metaplasia consistent with indeterminate for UIP.

724 Tissue from patient F9, II.1 showed diffuse uniform fibrosis without subpleural or paraseptal
725 predominance and without fibroblast foci and was classified as indeterminate for UIP.

726 Tissue from patient F14, II.1 showed diffuse fibrosis with microscopic honeycombing and
727 without persistence of unaffected lung. Prominent pneumocyte hyperplasia was observed. This
728 case was classified as indeterminate for UIP.

729 SP-A immunostaining (right column, x25) documented SP-A expression by hyperplastic
730 pneumocytes without a continuous secreted SP-A linear layer observed in the control individual
731 (insert X200).

732

733 **Figure 5: High-resolution computed tomography (CT) scans of the probands carrying the**
734 ***SFTPA1* / *SFTPA2* mutations.**

735 The CT-scans were separated into A: predominant ground glass opacities with or without signs
736 of fibrosis; B: lung fibrosis without predominant ground-glass opacities. Other features of ILD
737 were basal thickened septa, micronodules, reticulations, consolidations, honeycombing, and
738 bronchiectasis. The proposed classification is provided for each patient in **Table 2** and
739 **Supplemental Table 1** and includes usual interstitial pneumonia (UIP), non-specific interstitial

740 pneumonia (NSIP), pleuroparenchymal fibroelastosis (PPFE), hypersensitivity pneumonitis
741 (HSP), indeterminate pattern

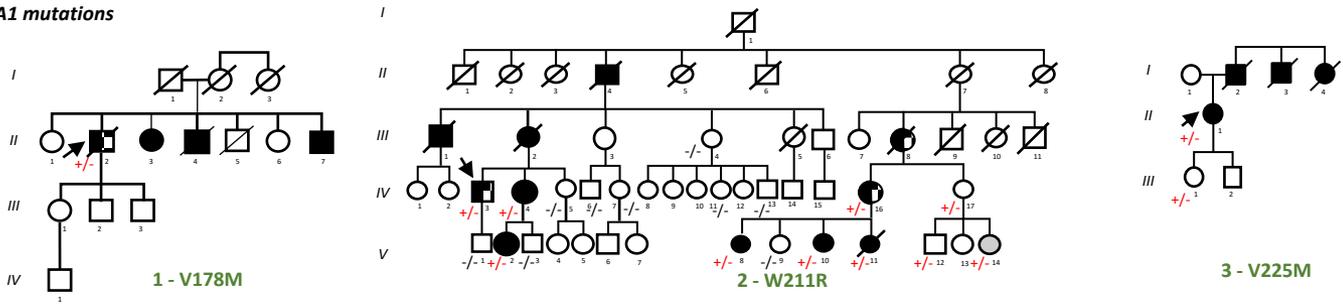
742

743 **Figure 6: Kaplan-Meier survival curves with or without transplantation; and post-**
744 **transplantation.**

745 **A:** Survival rate without (n=17) and with (n=11) transplantation of the 28 individuals with
746 identified *SFTPA1/A2* mutations. The age at diagnosis / age at screening was used as a starting
747 point. **B.** Survival rate after transplantation of the 11 patients with identified *SFTPA1/A2*
748 mutations. **A and B:** As all the patients did not complete the same duration of follow-up, the
749 end of follow-up of each patient is represented by a point.

750

SFTPA1 mutations



	ILD
	Lung cancer
	ILD and lung cancer
	Undefined chronic lung disease

SFTPA2 mutations

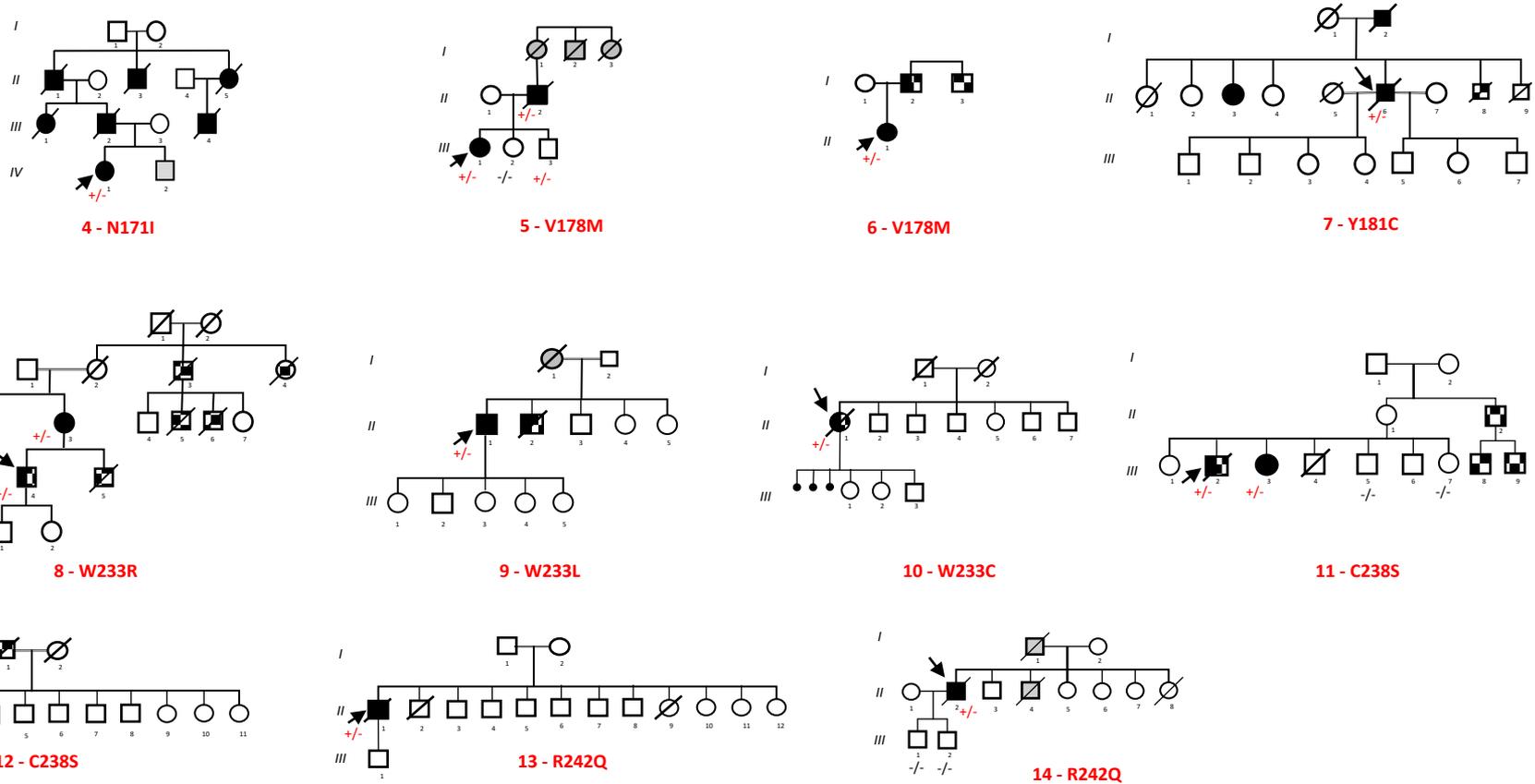


Figure 2: Genealogical trees of the 14 included families with *SFTP1* or *SFTP2* mutations

The arrows indicate the probands. Individuals with interstitial lung disease (ILD) are indicated by a black symbol. Individuals with lung cancer are indicated by a checkerboard symbol. Symbols that are half black and half checkerboard indicate individuals with ILD and lung cancer. Light grey symbols represent undefined chronic lung diseases reported by the family. Each individual is identified by his generation (Roman numbers) and its number (Arabic number). +/- indicates the presence of a *SFTP1* or *SFTP2* mutation in the heterozygous state; -/- indicates an absence of *SFTP1* or *SFTP2* mutation.

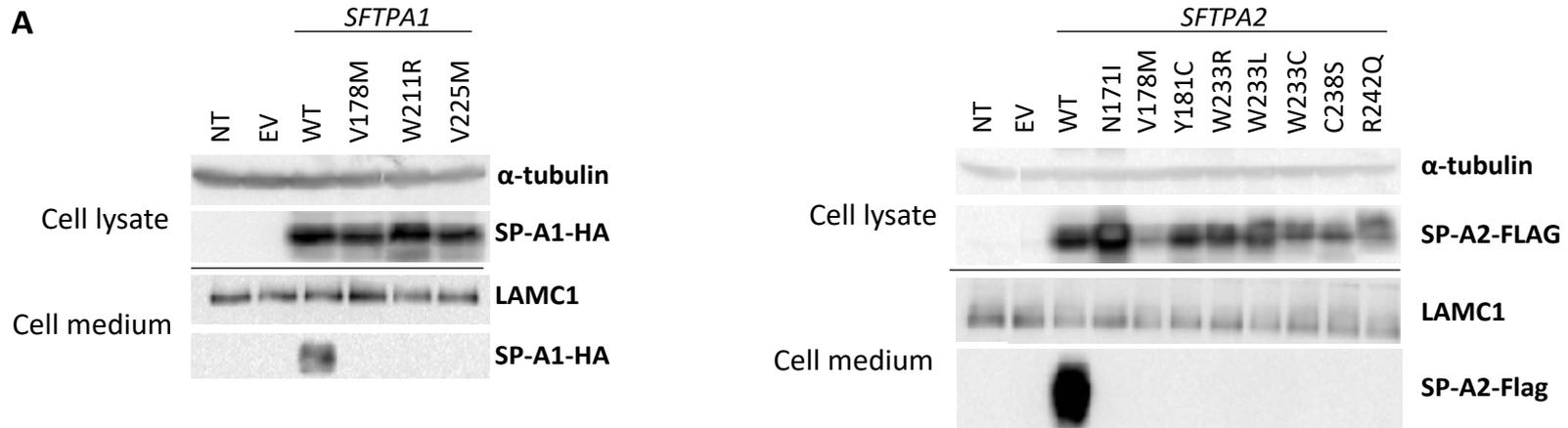
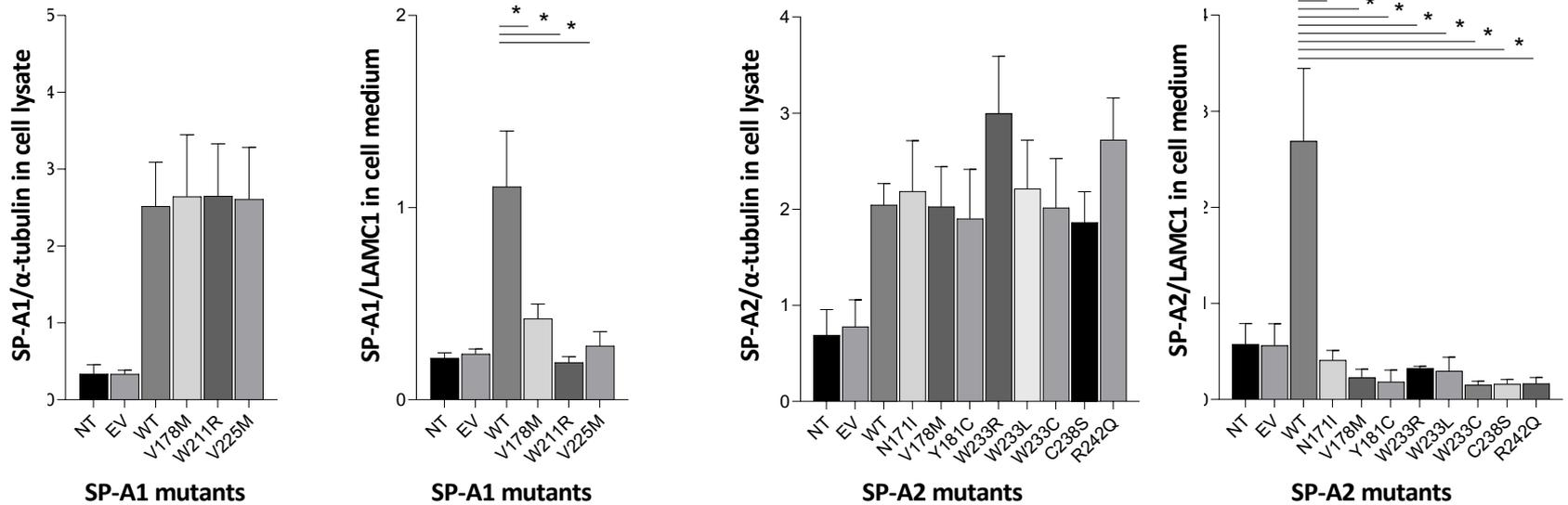
A**B**

Figure 3: Effects of *SFTP A1* and *SFTP A2* variants on protein production and secretion in HEK293T cells

A. Western blot. A 35-37-kDa molecular species consistent with the HA-tagged SP-A1 or FLAG-tagged SP-A2 proteins is observed in lysates of cells expressing the wild-type (WT) or mutant SP-A isoforms. A similar 35-37-kDa molecular species is also observed in the medium of cells expressing the WT proteins, but in the medium of the cells transfected with the mutant constructs of SP-A1 or SP-A2, an absence of secretion is observed. α -tubulin production was used as a loading control of protein expression and LAMC1 secretion was used as a loading control of protein secretion.

B. Normalized protein expression based on densitometry from western blot experiments. Charts show HA-tagged SP-A1 and FLAG-tagged SP-A2 protein expression normalized by α -tubulin (cell lysate expression) and LAMC1 (cell medium expression). Data are means of densitometry ratio (Image J software) from 3 or more independent western blot experiments. Error bars show standard error of the mean.

NT: not transfected; WT: wild type; EV: empty vector; *, $p < 0.05$

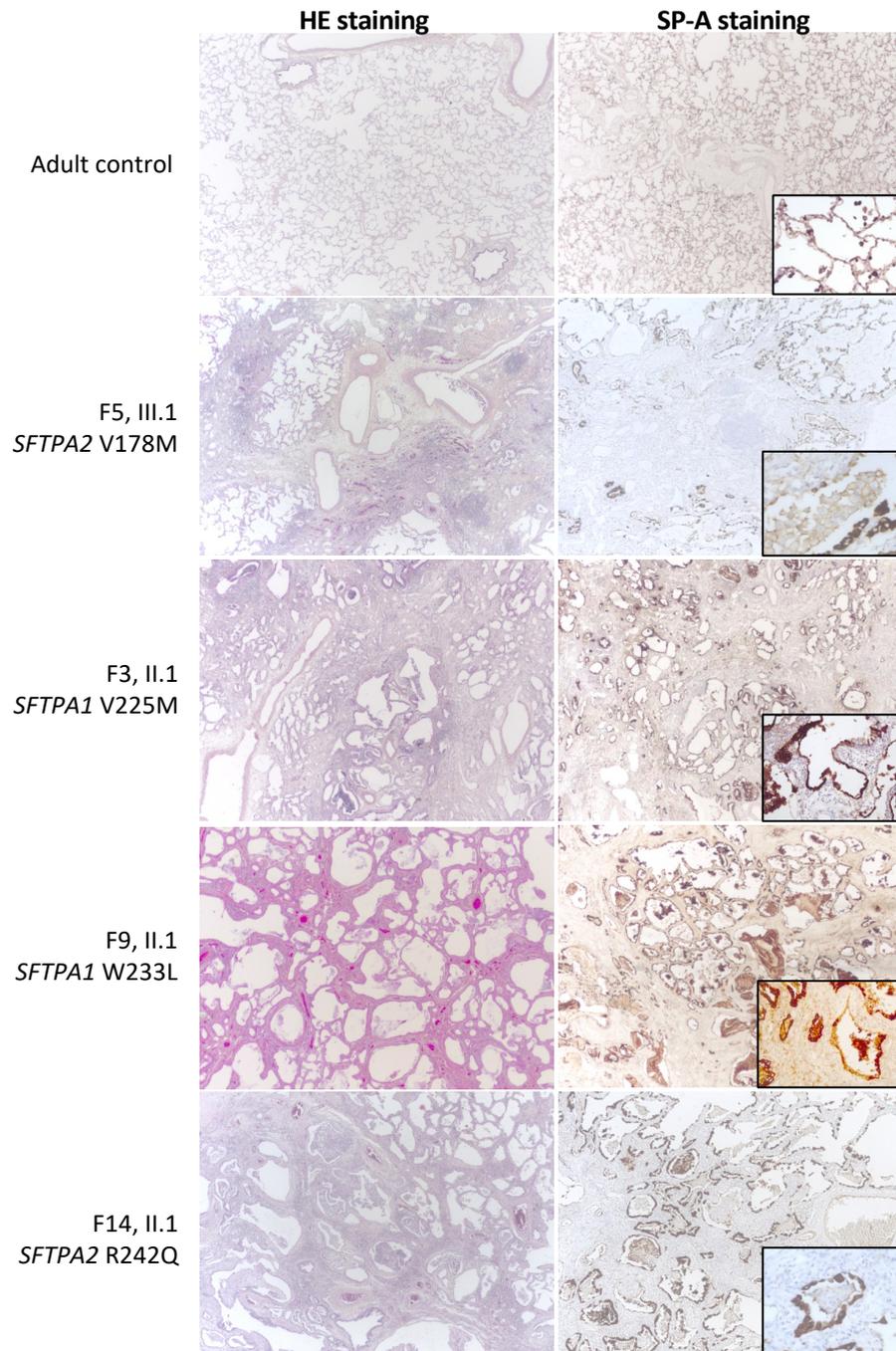


Figure 4: Lung histopathological features and SP-A expression on lung tissue from patients heterozygous for a *SFTP1* or *SFTP2* mutation

Histological examination was performed on lung tissues from 4 index patients (family 5, III.1; family 3, II.1; family 14, II.1; and family 9, II.1) and compared to an adult lung tissue sample. Haematoxylin and eosin (HE) staining was performed and SP-A expression was analyzed by immunohistochemistry. In all the patients, the HE staining (left column, x25) confirmed the ILD, with however various patterns described according to the 2018 ATS/ERS/JRS/ALAT guideline [3].

Tissue from patient F5, III.1 shows diffuse fibrosis without fibroblastic foci delimitating some unaffected lung parenchyma associated with prominent inflammation with lymphoid nodules consistent with indeterminate for UIP.

Tissue from patient F3, II.1 shows diffuse fibrosis with very few unaffected lung parenchyma with microscopic honeycombing, inflammation away from areas of honeycombing, bronchiolar metaplasia and osseous metaplasia consistent with indeterminate for UIP.

Tissue from patient F9, II.1 shows diffuse uniform fibrosis without subpleural or paraseptal predominance and without fibroblast foci and was classified as indeterminate for UIP.

Tissue from patient F14, II.1 shows diffuse fibrosis with microscopic honeycombing and without persistence of unaffected lung. Prominent pneumocyte hyperplasia was observed. This case was classified as indeterminate for UIP.

SP-A immunostaining (right column, x25) documented SP-A expression by hyperplastic pneumocytes without a continuous secreted SP-A linear layer observed in the control individual (insert X200).

**A-Predominant GGO
(with or without fibrosis)**



F6, II.1, 34 yo, *SFTPA2* V178M



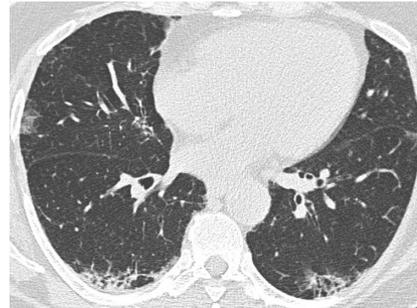
F2, IV.3, 45 yo, *SFTPA1* W211R



F8, IV.4, 34 yo, *SFTPA2* W233R



F9, II.1, 45 yo, *SFTPA2* W233L



F10, II.1, 50 yo, *SFTPA2* W233C

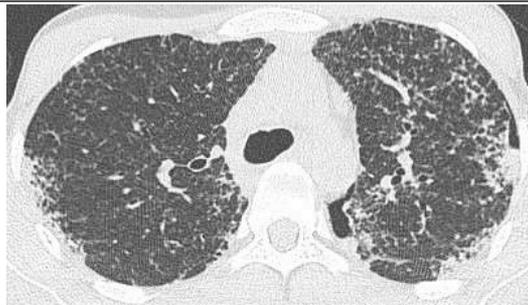


F12, II.1, 26 yo, *SFTPA2* C238S

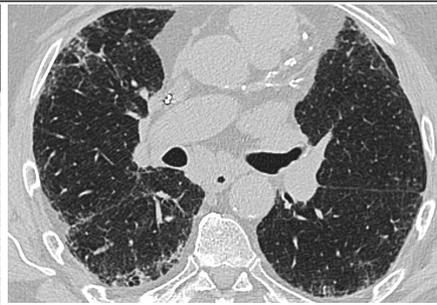


F14, II.1, 51 yo, *SFTPA2* R242Q

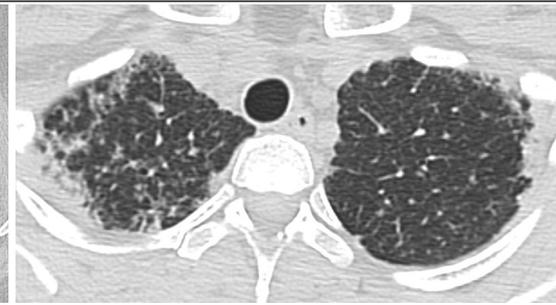
**B- Lung fibrosis
without
predominant GGO**



F4, IV.1, 32 yo, *SFTPA2* N171I



F1, II.2, 64 yo, *SFTPA1* V178M



F5, III.1, 36 yo, *SFTPA2* V178M

Figure 5: High-resolution computed tomography (CT) scans of the probands carrying the *SFTP1* / *SFTP2* mutations.

The CT-scans were separated into A: predominant ground glass opacities with or without signs of fibrosis; B: lung fibrosis without predominant ground-glass opacities. Other features of ILD were basal thickened septa, micronodules, reticulations, consolidations, honeycombing, and bronchiectasis. The proposed classification is provided for each patient in **Table 2** and **Supplemental Table 1** and includes usual interstitial pneumonia (UIP), non-specific interstitial pneumonia (NSIP), pleuroparenchymal fibroelastosis (PPFE), hypersensitivity pneumonitis (HSP), indeterminate pattern.

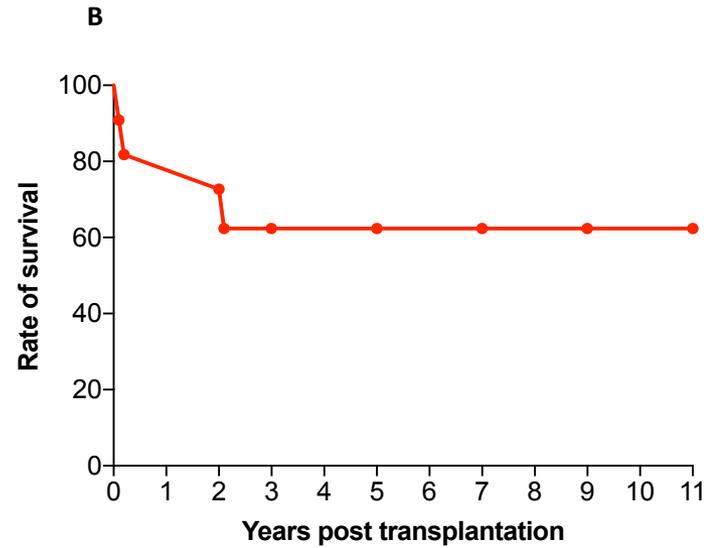
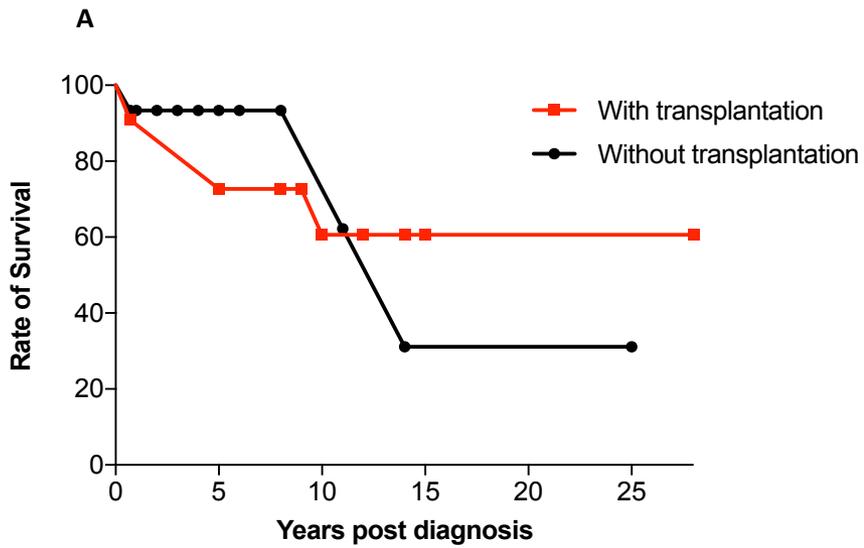


Figure 6: Kaplan-Meier survival curves with or without transplantation and post-transplantation.

A. Survival rate without (n=17) and with (n=11) transplantation of the 28 individuals with identified *SFTPA1/A2* mutations. The age at diagnosis / age at screening was used as a starting point. **B.** Survival rate after transplantation of the 11 patients with identified *SFTPA1/A2* mutations. **A and B:** As all the patients did not complete the same duration of follow-up, the end of follow-up of each patient is represented by a point.