AP-HP.Sorbonne Université
Charles-Foix • Pitié-Salpêtrière
Rothschild • Saint-Antoine • Tenon
Trousseau • La Roche-Guyon

Département Médico-Universitaire de Biologie et Génomique Médicales BioGeM
Directrice Médicale : Pr. Rachel LEVY

HOPITAL ARMAND TROUSSEAU
26 avenue du Dr Arnold Netter
75571 PARIS Cedex 12

Département de Génétique Médicale
Responsable : Pr. Jean-Pierre SIFFROI

Cadre de santé médicotechnique :
Mme Hélène MONTRESOR
☎ : 01 44 73 52 74

U.F. de Génétique Moléculaire
Responsable : Pr. Serge AMSELEM
☎ : 01 44 73 52 95
Fax : 01 44 73 52 19

Secteur Ketty Schwartz

Secrétariat :
Annick SALLES
☎ : 01 44 73 52 95
Fax : 01 44 73 52 19

Dr. Irina GIURGEA (MCU-PH)
☎ : 01 44 73 52 45

Dr. Grégory JOUVION (MCU-PH)
☎ : 01 71 73 88 18

Dr. Marie LEGENDRE (PH)
☎ : 01 44 73 52 21

E-mail : prenom.nom@aphp.fr

ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS

Membres du COPIL Recherche COVID

Paris, le 15 avril 2019

Chers Collègues,

Nous sollicitons le COPIL Recherche COVID concernant notre projet qui porte sur l'identification de facteurs génétiques de prédisposition aux formes sévères de COVID-19 et de potentielles cibles thérapeutiques.

La maladie COVID-19 est caractérisée par un large spectre phénotypique allant des formes extrêmement sévères aux formes asymptomatiques. La description de formes sévères chez de jeunes patients, et inversement des formes paucisymptomatiques chez des sujets plus âgés, suggère fortement l'existence des facteurs d'hôte qui conditionneraient la sévérité de la maladie. Cette question nous interpelle tout particulièrement pour deux raisons essentielles :

(i) le phénotype des formes sévères caractérisées par un état hyperinflammatoire et un « orage cytokinique » rappelle celui observé dans les maladies auto-inflammatoires (MAI) et la lymphohistiocytose hémophagocytaire pour lesquelles plusieurs gènes responsables ont déjà été identifiés.

(ii) nous travaillons de très longue date sur l'étude des bases moléculaires et cellulaires des MAI, pour lesquelles, en tant que centre référent national, nous assurons le diagnostic au sein de **l'UF de génétique moléculaire** de l'hôpital Trousseau (dir. S. Amselem) (> 10.000 ADN testés pour des gènes dont nous connaissons très bien la variabilité moléculaire). Sur un plan plus fondamental, nous étudions la physiopathologie des MAI au sein de l'unité de recherche **UMR_S933 Inserm/SU** (dir. S. Amselem) attenante au secteur diagnostique, avec notamment l'évaluation, sur différents modèles cellulaires maîtrisés, des conséquences fonctionnelles des variations moléculaires identifiées.

Pour cette étude pilote qui est financée par la Faculté de médecine Sorbonne Université, nous avons donc **la capacité de tester très rapidement** (en quelques semaines) si des variations de séquence dans ces excellents gènes candidats – auxquels nous ajoutons les gènes codant des protéines clés pour l'entrée du virus dans les cellules de l'hôte—, contribuent à la sévérité de COVID-19. En effet, mes responsabilités à la fois dans le diagnostic moléculaire hospitalier et les études plus fondamentales des MAI me permettent de mener ce type d'étude. Notre **approche NGS ciblée** (panel d'une centaine de gènes) offre la possibilité d'identifier les **mutations somatiques en mosaïque, ce qui est déterminant au vu du nombre de patients âgés atteints de COVID-19** (Louvrier C,… Giurgea I, *J Allergy Clin Immunol* 2020 ; Assrawi E,…Giurgea I, *J Invest Dermatol* 2020). Une approche par whole exome (WES) ou whole genome (WGS) peut être envisagée secondairement, d'autant que le WES est réalisé en routine dans notre laboratoire diagnostique et que 450 patients avec MAI seront prochainement étudiés par WGS par notre équipe dans le cadre du programme européen ImmunAID (H2020).

Nous vous remercions de l'attention que vous porterez à notre demande d'accès à la cohorte COVIDeF et restons à votre disposition pour toute information complémentaire.

Dr. Irina Giurgea

Severe forms of COVID-19: Genetic predisposing factors and therapeutic issues

Irina Giurgea - UMR933 Inserm/Sorbonne Université_Unité Fonctionnelle de génétique moléculaire (APHP Hôpital Trousseau)

RÉSUMÉ DU PROJET

The current pandemic infectious disease COVID-19, which is due to the SARS-CoV-2 (severe acute respiratory syndrome-coronavirus 2) virus, causes significant morbidity and mortality. COVID-19 is, however, characterized by a very large phenotypic spectrum ranging from extremely severe forms manifesting as an acute respiratory distress syndrome (ARDS) to asymptomatic forms. On the one hand, key observations strongly suggest that host factors of genetic origin might contribute to the variable phenotypic expression of COVID-19 among infected individuals. On the other hand, patients with a very severe form of COVID-19 display a cytokine storm syndrome resembling haemophagocytic lymphohistiocytosis (HLH) in which several causative genes have been identified. The hyperinflammatory state reported in COVID-19 and in HLH is also highly reminiscent of the one observed in systemic autoinflammatory diseases (SAIDs) that are also explained by gene mutations. We hypothesize that the genes encoding the proteins belonging to the pathways involved in the pathogenesis of HLH, SAIDs as well as inflammatory interstitial lung diseases (ILDs) are excellent candidates to account for the extreme phenotypic variability of COVID-19. Indeed, all these genes encode a wide variety of key components of the innate immune system that is known to be involved in the antiviral defence of virus-naïve patients. The major theme of our research unit is dedicated to the study of innate immune factors (genes/proteins) involved in SAIDs and ILDs with a strong inflammatory component. Therefore, the aim of this pilot study, which relies on our expertise in the genetics of autoinflammation, is to rapidly test whether sequence variations in a series of candidate genes including the genes involved in HLF, SAIDs as well as those involved in virus entry into host cells, can contribute to COVID-19 severity. Identification of such genetic factors of prognostic value could be of particular help to greatly improve disease management and to unveil actionable therapeutic targets.

RATIONNEL

The current pandemic infectious disease COVID-19, which is due to the SARS-CoV-2 (severe acute respiratory syndrome-coronavirus 2) virus, causes significant morbidity and mortality (by 26 March, more than 21,000 deaths have been attributed to COVID). **COVID-19 is however characterized by a very large phenotypic spectrum ranging from extremely severe forms manifesting as an acute respiratory distress syndrome (ARDS) to asymptomatic forms.** Elderly patients (≥70 years old) and people with underlying diseases such as respiratory diseases, cardiovascular diseases, diabetes or immune deficits are known to be at risk of developing a severe form of the disease. However, more rarely severe forms of COVID-19 can also occur in young people (≤50 years old) and, conversely, elder patients can also present a mild form of the disease[1] suggesting that **genetic factors (modifier genes) may influence disease progression**. The phenotypic spectrum of COVID-19 can also be schematically categorized according to the patients' disease outcome (those seen at the Emergency Department who will recover; those requiring hospitalization but no intensive care; and those requiring intensive care). The current project is built on these key observations that strongly suggest **the contribution of host factors of genetic origin to the variable phenotypic expression of COVID-19 among infected individuals.**

1/ Phenotypic and molecular aspects

Patients with a very severe form of COVID-19 display a **cytokine storm syndrome** resembling what is seen in **haemophagocytic lymphohistiocytosis (HLH)**[2], which is a **hyperinflammatory syndrome** mediated by immune cells including persistently activated macrophages that derive from monocytes. HLH symptomatology includes unremitting fever, cytopenias, and hyperferritinaemia; most importantly, pulmonary involvement (including ARDS) occurs in approximately 50% of HLH patients[3]. Strikingly, a HLH-resembling cytokine profile has been observed in COVID-19 (i.e. high plasma levels of interleukin (IL)-2, IL-7, granulocyte colony stimulating factor, interferon-γ inducible protein 10, monocyte chemoattractant protein 1, macrophage inflammatory protein 1-α, and tumor necrosis factor-α[1]).

The hyperinflammatory state present in COVID-19 and in HLH is also highly reminiscent of the one observed in the so-called "**systemic autoinflammatory diseases" (SAIDs)**. SAIDs, which are characterised by extensive clinical and biological inflammation, consist of an evolving group of conditions: (i) the inflammasomopathies (involving multiprotein complexes named infammasomes such as the pyrin inflammasome involved in familial Mediterranean fever or FMF and NLRP3 and NLRC4 inflammasomes in NLRP3- and NLRC4- autoinflammatory disorders, respectively)[4,5], and (ii) the interferonopathies, associated with type I interferon (IFN) upregulation. This is the case of the **interstitial lung disease (ILD)** due to activating mutations in *TMEM173* (transmembrane protein 173), a gene that encodes STING (**st**imulator of **in**terferon **g**enes). Noteworthy, this autosomal-dominant disease is characterized by systemic inflammation and interstitial lung disease[10] and *TMEM173* is expressed in type II pneumocytes, bronchial cells and alveolar macrophages[10]. So far, more than 40 genes have been shown to be involved in the pathophysiology of SAIDs[6], with both germline and somatic mosaic mutations identified in circulating blood cells[4,5]. The SAID genes encode for a wide variety of proteins, which are key components of the innate immune system. Noteworthy, in several SAIDs, HLH can occur and lead to a lethal issue[7,8]. From a therapeutic viewpoint, it is worth noting that depending on the disease gene identified and on the involved signalling pathway, specific biotherapies can be proposed to those patients. Indeed, IL1-inhibitors are highly efficient for the treatment of NLRP3-AID patients. JAK1/2 inhibitors are very helpful in several type I interferonopathies associated with increased IFN secretion[9].

2/ Working hypothesis

Overall, these data prompt us to hypothesize that genetic variations in *NLRP3, TMEM173* and genes involved in primary HLH (also called familial, FHL) and in other innate immune system factors involved in SAIDs may represent key factors predisposing to severe forms of COVID-19. Additional candidate genes involved in disease severity include those implicated in virus entry into host cells, such as the angiotensin-converting enzyme 2 (ACE2) receptor[14] and the transmembrane protease serine 2 (TMPRSS2)[15].

- Concerning the response of the host innate immune system to the viral infection, it is important to note that (i) the SARS-CoV papain-like protease (PLpro) inhibits STING/TBK1/IKKε-mediated activation of type I IFNs[12], and (ii) the SARS-CoV envelope (E) protein that displays an ion channel (IC) activity (E protein IC) activates the NLRP3 inflammasome (through Ca^{2+} transport modification)[13].
- As for NLRP3 inflammasome deregulation, we aim to look for germinal and somatic mosaic *NLRP3* mutations in blood leukocytes, as routinely performed in our lab. Somatic mutations (mosaicism), could indeed readily account for the increased severity of COVID in elder patients.
- Common *TMEM173* alleles (frequency in the general population >1%) contain one to three non-synonymous single nucleotide polymorphisms (SNPs) such as the HAQ (R71H-G230A-R293Q) allele, which is the second most common allele of this gene. Studies using B-cell lines derived from

homozygous HAQ/HAQ individuals showed that HAQ/HAQ cells express extremely low level of the STING protein (encoded by *TMEM173*) and decreased *TMEM173* transcript suggesting that HAQ is a null allele[14]. Such relatively frequent allele is therefore of particular interest for the current study.

- The majority of FHL are autosomal recessive disorders (due to biallelic mutations in genes such as *PRF1*, *UNC13D*, *STX11*, *STXBP2*, *RAB27A, LYST, SH2D1A, AP3B1)* and one is X-linked (*XIAP*)[16]. The rationale to study these genes is based on the fact that heterozygous mutations in FLH genes are found in individuals with secondary HLH. In the majority of these patients a precipitating etiology was identified including infections[10–12]. Furthermore, the fact that men are more likely to be positive for SARS-CoV2 and to have a lethal form of COVID-19[17] prompt us to also focus on X-linked genes.
- Finally, the SARS-CoV-2 virus enters in the host cell through the binding of the viral Spike (S) glycoprotein to the host ACE2 receptor[14], expressed in the lungs, arteries, heart, kidneys, and the intestine. More explicitly, the S protein is cleaved into two subunits, S1 and S2, by an extracellular protease. While S1 binds to ACE2, S2 is further cleaved and activated by the host surface-associated TMPRSS2[15]. *ACE2*, which is located on the X chromosome, and *TMPRSS2* are therefore excellent candidate genes to be studied in COVID-19 patients.

3/ Team's assets for the current proposal

The major scientific theme of our research unit (UMR_S933) is dedicated to the study the pathophysiology of innate immune factors involved in SAIDs and ILDs with a strong inflammatory component. This research, which aims at identifying the molecular and cellular bases of these diseases, is tightly linked to our molecular diagnostics lab (UF de génétique moléculaire from the Département de Génétique – APHP.Sorbonne Université). This means that, on the one hand, we screen on a routine basis all the above-mentioned genes that are involved in SAIDs, HLH and inflammatory ILDs, and, on the other hand, we have developed molecular and cellular approaches to identify new causative genes in those conditions and to assess the functional consequences of the identified sequence variants.

Our contribution to the molecular bases of autoinflammation is the following: demonstration of the diagnostic value of the analysis of *MEFV,* the first SAID gene identified; this data validated the first objective criterion for the diagnosis of FMF[20]; demonstration of the prognostic value of *MEFV* molecular analysis, by establishing the association of a particular genotype (M694V homozygous), at risk for the development of renal amyloidosis[21]; identification of the first modifier gene for the FMF phenotype (SAA1), of which a particular genotype (alpha/alpha) is associated with an increased risk of developing inflammatory amyloidosis (OR=6.9)[21]; identification of two SAID genes (*NLRP12* and *TNFRSF11A*)[22,23]; demonstration of the existence of a FMF-like condition unrelated to *MEFV*[13]; identification of 14-3-3 proteins as Pyrin partners[24], with key implications for the understanding of Pyrin function and the pathophysiology of FMF and other *MEFV*-related auto-inflammatory disorders[25]; identification of the phenotypic and molecular characteristics of germline vs. somatic mosaic *NLRP3* mutations in autoinflammation[4]. The lab has also identified a gene (*SFTPA1*) involved in the occurrence of ILD[26].

Overall, given the dramatic increase of the number of COVID-19 patients and the existence of very severe forms of the disease, there is an urgent need to identify biomarkers that could both predict the disease outcome and unveil new therapeutic targets. This could be of particular help to greatly improve disease management. On the basis of the above-mentioned context, it is tempting to speculate that genetic factors could contribute to the disease severity and that the genes encoding the proteins belonging to the pathways involved (i) in the pathogenesis of HLH, SAIDs and inflammatory ILDs and (ii) in virus entry into host cells could represent excellent candidates to be tested in the patients with COVID-19.

OBJECTIFS DU PROJET

Given the fact that COVID-19 is characterized by a very large phenotypic spectrum ranging from asymptomatic forms to extremely severe forms manifesting as ARDS, our objective is to rely on our expertise in the pathophysiology of auto-inflammatory conditions in order:

1/ To rapidly test whether sequence variations in a series of candidate genes can contribute to COVID-19 severity (i.e. genes involved in SAIDs including *NLRP3*, in HLH and in inflammatory ILDs, as well as genes involved in virus entry into host cells).

2/ To subsequently analyse the whole exome of patients with no mutation identified in those preselected candidate genes.

MÉTHODOLOGIE & MISE EN ŒUVRE

1/ We will work on genomic DNA extracted from peripheral blood leukocytes to characterize the genetic profile of the patients. To this end, we plan to rely on the COVIDeF cohort of patients and its biobank currently set up by the Assistance Publique-Hôpitaux de Paris (Head: Pierre Hausfater). In this pilot study, we will first focus on patients belonging to the following 3 groups:

- Group I: Pauci-symptomatic patients seen at the Emergency Department and who will recover (and, if available, asymptomatic individuals/relatives). **In this group, the molecular data obtained in patients ≥70 years should unveil "protective" genetic factors.**

- Group II: Patients with a moderate form of the disease requiring hospitalization but no intensive care.

- Group III: Patients requiring intensive care with a severe form of the disease (i.e. ARDS and other manifestations that will be considered as severe in the above-mentioned cohort) and with no known co-morbidity. **In this group, the molecular data obtained in the patients ≤50 years should unveil "aggravating" genetic factors.**

2/ Molecular analyses will be performed through the following two approaches.

a. Candidate-gene approach: A targeted sequencing approach will be performed by next-generation sequencing (NGS) on a NextSeq500 or MiSeq (Illumina) available in the lab, using a custom targeted capture (SeqCap EZ Choice system; Roche) of the exons and the flanking intronic sequences of (i) the ~40 known SAID- and HLH-causing genes, including *NLRP3* and *TMEM173,* and (ii) ~100 additional candidate genes for inflammatory diseases, which include genes playing a potential role in the viral propagation and the host response to infection. Such a gene panel is already used in the lab on a routine basis, especially for the study of SAIDs. Somatic mutations (mosaicism) will also be sought thanks to the high-depth sequencing of the gene panel (mean sequence coverage per base 500X), as already performed in our team[4,5].

b. Whole exome sequencing: WES will be performed in the patients with no mutation identified after targeted sequencing of the genes present on our panel. This will be done using the Nextseq500 (Illumina) that is available in the lab. NGS data will be analyzed as previously reported by our team[4].

3/ We plan to start the study as soon as we have access to the COVID-19 cohort, ideally with a minimum of 50 patients per group. Data analyses will be performed within each group and will be compared between the 3 groups.

4/ Whatever the approach followed in order to discover gene variants potentially involved in COVID-19, depending on these results, the gene involved and the nature of the mutation identified, we will

perform specific functional studies that are not part of the current application. Briefly, these functional studies will rely on the cellular models we have set up in the lab to assess the consequences of the sequence variants we identify in our candidate genes[5,27,28].

RÉSULTATS ATTENDUS

The ongoing set up of COVID-19 cohort together with the long-standing expertise of our research unit in the pathophysiology of both SAIDs and inflammatory ILDs provide a unique opportunity to foster the discovery of biomarkers predictive of the severity of COVID-19. **Overall, the data obtained could help i) providing a rapid and accurate prognostic tool, ii) deciphering the biological networks involved in disease severity, and therefore iii) identifying new potential therapeutic targets.**

ÉQUIPES IMPLIQUÉES

Équipe pilote : Dr. Irina Giurgea, UMR_S933 Inserm/Sorbonne Université (Dir S. Amselem).
Équipe(s) partenaire(s) : The methodological analyses will be performed by Pr. Florence Tubach, Département BIOSPIM and Unité de recherche Clinique, APHP.Sorbonne Université.

BUDGET

This project is supported by Sorbonne Université (69 000 €) in response to the COVID-19 call launched by the Faculté de Médecine Sorbonne Université, and by our research unit UMR_S933.

RÉFÉRENCES

* Publications from UMR_S933 team members

1 Huang C, Wang Y, Li X *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; **395**: 497–506.

2 Mehta P, McAuley DF, Brown M *et al.* COVID-19: consider cytokine storm syndromes and immunosuppression. *Lancet* 2020. doi:10.1016/S0140-6736(20)30628-0.

3 Seguin A, Galicier L, Boutboul D, Lemiale V, Azoulay E. Pulmonary Involvement in Patients With Hemophagocytic Lymphohistiocytosis. *Chest* 2016; **149**: 1294–1301.

4* Louvrier C, Assrawi E, El Khouri E *et al.* NLRP3-associated autoinflammatory diseases: Phenotypic and molecular characteristics of germline versus somatic mutations. *J Allergy Clin Immunol* 2019. doi:10.1016/j.jaci.2019.11.035.

5* Assrawi E, Louvrier C, Lepelletier C *et al.* Somatic Mosaic NLRP3 Mutations and Inflammasome Activation in Late-Onset Chronic Urticaria. *J Invest Dermatol* 2020; **140**: 791-798.e2.

6 de Jesus AA, Canna SW, Liu Y, Goldbach-Mansky R. Molecular mechanisms in genetically defined autoinflammatory diseases: disorders of amplified danger signaling. *Annu Rev Immunol* 2015; **33**: 823–874.

7 Canna SW, de Jesus AA, Gouni S *et al.* An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. *Nat Genet* 2014; **46**: 1140–1146.

8 Kitamura A, Sasaki Y, Abe T, Kano H, Yasutomo K. An inherited mutation in NLRC4 causes autoinflammation in human and mice. *J Exp Med* 2014; **211**: 2385–2396.

9 Sanchez GAM, Reinhardt A, Ramsey S *et al.* JAK1/2 inhibition with baricitinib in the treatment of autoinflammatory interferonopathies. *J Clin Invest* 2018; **128**: 3041–3052.

10 Liu Y, Jesus AA, Marrero B *et al.* Activated STING in a vascular and pulmonary syndrome. *N Engl J Med* 2014; **371**: 507–518.

11 Ishikawa H, Barber GN. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* 2008; **455**: 674–678.

12 Chen X, Yang X, Zheng Y, Yang Y, Xing Y, Chen Z. SARS coronavirus papain-like protease inhibits the type I interferon signaling pathway through interaction with the STING-TRAF3-TBK1 complex. *Protein Cell* 2014; **5**: 369–381.

13 Nieto-Torres JL, Verdiá-Báguena C, Jimenez-Guardeño JM *et al.* Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. *Virology* 2015; **485**: 330–339.

14 Zhou P, Yang X-L, Wang X-G *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020; **579**: 270–273.
15 Hoffmann M, Kleine-Weber H, Schroeder S *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020. doi:10.1016/j.cell.2020.02.052.
16 Cetica V, Sieni E, Pende D *et al.* Genetic predisposition to hemophagocytic lymphohistiocytosis: Report on 500 patients from the Italian registry. *J Allergy Clin Immunol* 2016; **137**: 188-196.e4.
17 Wu C, Chen X, Cai Y *et al.* Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med* 2020. doi:10.1001/jamainternmed.2020.0994.
18 de Wit E, van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: recent insights into emerging coronaviruses. *Nat Rev Microbiol* 2016; **14**: 523–534.
19 Davis ME, Gack MU. Ubiquitination in the antiviral immune response. *Virology* 2015; **479–480**: 52–65.
20* Cazeneuve C, Sarkisian T, Pecheux C *et al.* MEFV-Gene analysis in armenian patients with Familial Mediterranean fever: diagnostic value and unfavorable renal prognosis of the M694V homozygous genotype-genetic and therapeutic implications. *Am J Hum Genet* 1999; **65**: 88–97.
21* Cazeneuve C, Ajrapetyan H, Papin S *et al.* Identification of MEFV-independent modifying genetic factors for familial Mediterranean fever. *Am J Hum Genet* 2000; **67**: 1136–1143.
22* Jeru I, Duquesnoy P, Fernandes-Alnemri T *et al.* Mutations in NALP12 cause hereditary periodic fever syndromes. *Proc Natl Acad Sci U S A* 2008; **105**: 1614–9.
23* Jéru I, Cochet E, Duquesnoy P *et al.* Brief Report: Involvement of TNFRSF11A molecular defects in autoinflammatory disorders. *Arthritis & Rheumatology (Hoboken, NJ)* 2014; **66**: 2621–2627.
24* Jeru I, Papin S, L'hoste S *et al.* Interaction of pyrin with 14.3.3 in an isoform-specific and phosphorylation-dependent manner regulates its translocation to the nucleus. *Arthritis Rheum* 2005; **52**: 1848–1857.
25* Masters SL, Lagou V, Jéru I *et al.* Familial autoinflammation with neutrophilic dermatosis reveals a regulatory mechanism of pyrin activation. *Sci Transl Med* 2016; **8**: 332ra45.
26* Nathan N, Giraud V, Picard C *et al.* Germline SFTPA1 mutation in familial idiopathic interstitial pneumonia and lung cancer. *Hum Mol Genet* 2016; **25**: 1457–1467.