



**HAL**  
open science

## **New surfactant protein C gene mutations associated with diffuse lung disease**

Loïc Guillot, Ralph Epaud, Guillaume Thouvenin, Laurence Jonard, Amira Mohsni, Remy Couderc, François Counil, Jacques de Blic, Rola Abou Taam, Muriel Le Bourgeois, et al.

► **To cite this version:**

Loïc Guillot, Ralph Epaud, Guillaume Thouvenin, Laurence Jonard, Amira Mohsni, et al.. New surfactant protein C gene mutations associated with diffuse lung disease. *Journal of Medical Genetics*, 2009, 46 (7), pp.490. 10.1136/jmg.2009.066829 . hal-00552694

**HAL Id: hal-00552694**

**<https://hal.science/hal-00552694>**

Submitted on 6 Jan 2011

**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.

## **New surfactant protein C gene mutations associated with diffuse lung disease**

Loïc Guillot<sup>1,2 †</sup>, Ralph Epaud<sup>1,2,3 †</sup>, Guillaume Thouvenin<sup>1,2</sup>, Laurence Jonard<sup>4</sup>, Amira Mohsni<sup>4</sup>, Remy Couderc<sup>4</sup>, François Counil<sup>5</sup>, Jacques de Blic<sup>6</sup>, Rola Abou Taam<sup>6</sup>, Muriel Le Bourgeois<sup>6</sup>, Philippe Reix<sup>7</sup>, Florence Flamein<sup>1,2</sup>, Annick Clement<sup>1,2,3</sup>, Delphine Feldmann<sup>4\*</sup>.

<sup>1</sup>INSERM UMR\_S U938, Paris, France; <sup>2</sup>UPMC Univ Paris 06, France; <sup>3</sup>AP-HP, Hôpital Armand Trousseau, Pediatric Pulmonary Department, Paris, France; <sup>4</sup>AP-HP, Hôpital Armand Trousseau, Biochemistry Department, Paris, France ; <sup>5</sup>Hôpital Arnaud de Villeneuve, Centre Hospitalier Universitaire de Montpellier, Montpellier, France; <sup>6</sup>Université Paris Descartes, AP-HP, Hôpital Necker Enfants Malades, Pediatric Pneumology-allergology Department, Paris, France; <sup>7</sup>Groupement Hospitalier Est, Pediatric Pneumology-Allergology Department, Lyon, France. <sup>†</sup>These authors equally contributed to this work.

\*Correspondence to Delphine Feldmann

Hôpital Armand Trousseau

Laboratoire de Biochimie

26, Avenue du Dr Arnold Netter

75571 Paris cedex 12, France

Email: delphine.feldmann@trs.aphp.fr

Phone: +33 144736867

Fax: +33 1 44736687

Running Title: New *SFTPC* mutations associated with lung disease.

## ABSTRACT

Mutations in the surfactant protein C gene (*SFTPC*) have been recently associated with the development of diffuse lung disease, particularly sporadic and familial interstitial lung disease (ILD). We have investigated the prevalence and the spectrum of *SFTPC* mutations in a large cohort of infants and children with diffuse lung disease and suspected with surfactant dysfunction. One hundred twenty-one children were first screened for the common *SFTPC* mutation, p.Ile73Thr (I73T). Ten unrelated patients were shown to carry this mutation. The I73T mutation was inherited in 6 cases, and appeared *de novo* in 4. The 111 patients without the I73T mutation were screened for the entire coding sequence of *SFTPC*. Of these, eight (seven unrelated) subjects were shown to carry a novel mutant allele of *SFTPC*. All these seven new mutations are located in the BRICHOS domain except the p.Val39Ala (V39A) mutation, which is in the surfactant protein C (SP-C) mature peptide. Our results confirm that *SFTPC* mutations are a frequent cause of diffuse lung disease, and that I73T is the most frequent *SFTPC* mutation associated with diffuse lung disease.

KEY WORDS: children, lung, surfactant protein C (SP-C), *SFTPC*

## INTRODUCTION

Diffuse lung disease in children, including interstitial lung disease (ILD) and chronic lung disease, is a rare heterogeneous group of chronic disorders characterized by impaired gas exchange<sup>1</sup>. Recent studies suggest that surfactant deficiency plays a role in the pathogenesis<sup>2</sup>. Pulmonary surfactant is a complex mixture of proteins and lipids that prevents collapse at end expiration by reducing surface tension at the air-water interface of the lung alveoli.

Surfactant protein C (SP-C), a hydrophobic protein critically involved in surfactant homeostasis<sup>3</sup>, is expressed by alveolar type II epithelial cells and encoded by a single gene located on chromosome 8 (8p21.3) called *SFTPC*. Human *SFTPC* gene is organized into 6 exons (I through V coding, VI untranslated). SP-C is synthesized as a 197-amino acid proprotein (proSP-C) that undergoes multiple processing steps to the mature SP-C peptide of 35 amino acids, to be finally released in the alveoli associated with the other surfactant proteins and phospholipids<sup>3</sup>. Nogee et al. first described a *SFTPC* mutation in 2001<sup>4</sup>. The mutation, c.435+1G>A, located at the first base of intron 4 alters the normal donor splice site and results in the skipping of exon 4 and the deletion of 37 amino acids in the carboxy-terminal domain of proSP-C. This mutation was observed in one allele of a patient suffering from chronic interstitial pneumonitis and was inherited following a dominant pattern. To date, 13 *SFTPC* mutations have been identified<sup>4-11</sup>. They always occur in one single allele and have been associated with diffuse lung disease in children and adults. The mutation I73T (c.218 T>C) is the more prevalent mutation; others have been described in only one family. Dominant transmission with variable penetrance has been confirmed in familial cases, but a number of sporadic cases have also been reported with a *de novo* mutation. The phenotype associated with *SFTPC* mutations appears to be very variable. Indeed, neonatal forms leading to death in the

first years of life as well as childhood and adult forms with chronic respiratory disease have been reported. Single nucleotide polymorphisms (SNPs) have also been detected both in the coding and non-coding sequence, with unclear consequences on proSP-C processing and function <sup>2</sup>.

The variable phenotype with unspecific biological findings makes the genetic diagnosis of SP-C inherited disease essential. Thus, the present study aims to investigate *SFTPC* variations in a large population of 121 children (120 unrelated) with diffuse lung disease and suspected surfactant dysfunction.

## **MATERIALS AND METHODS**

### *Patients*

One hundred and twenty-one newborns/children (60 male, 61 female) diagnosed with diffuse lung disease were recruited over a 5-year period (2002–2007) through a national program on rare lung disease by several university hospital-based pediatric pulmonology practices (Hôpital Armand Trousseau, Paris; Hôpital Necker Enfants Malades, Paris; Groupement Hospitalier Est, Lyon and Centre Hospitalier Universitaire CHU Arnaud de Villeneuve, Montpellier). Seventy four patients originated from Europe, 24 from North Africa, 13 from Africa, 9 from Reunion Island and 1 from French Antilles. Among the 121 patients with diffuse lung disease, 86 had respiratory distress, 59 presenting with neonatal onset and 18 patients died. Clinical observations included family background, physical signs at presentation, chest radiography and computed tomography (CT), and histological lung analysis when available. The protocol was accepted by the Committee for the protection of individuals in

Biochemical research as required by French legislation and written informed consent was obtained from the patients or their next of kin included in this study.

The 121 patients where the common 121ins2 *SFTP*B mutation was initially ruled out were analyzed for the common I73T mutation. Then, the entire *SFTP*C coding sequence was performed in patients without the I73T mutation.

### *Mutation Analysis*

Genomic DNAs were extracted from patients' blood using an automated BioRobot EZ1 workstation (Qiagen, Hilden, Germany). The common mutation I73T was screened using an allele specific amplification as previously published<sup>12</sup> and confirmed by the direct sequencing of exon 3. The 5 coding exons of *SFTP*C and the exon-intron boundaries were analyzed by direct DNA sequencing of PCR products using the following primers: exon 1 sense 5'-ACCCAGGTTTGCTCTTGCT-3', exon 1 antisense 5'-TGAATGGATCTGGATAAGGAAA-3', exon 2 sense 5'-TGTTAGAATCCAGGCCACCT-3', exon 2 antisense 5'-CGTGCCTCTTTCCTTCTAGC-3', exon 3 sense 5'-CTCTTGGGAAAGAGGGAAGC-3', exon 3 antisense 5'-GGGAGAGATGGATGTGGATG-3', exon 4 sense 5'-CTAGTATGACTCCCGTGCCC-3', exon 4 antisense 5'-TGAGGAACAGTGCTTTACAGG-3', exon 5 sense 5'-TCAGCTGAGTCCACTCACTACC-3', exon 5 antisense 5'-GTACCGGTCTGTGAGCTTCC-3'. Sequencing reactions were performed with Big Dye Terminator cycle sequencing kits from Applied Biosystems (Foster city, CA) and run on an ABI Prism 3130 Genetic Analyzer. Sequences were analyzed using SeqScape software and compared with the reference sequence NM\_003018.3. Mutations are named following the international nomenclature, starting at nucleotide 1 of the first codon (Human Genome

Variation Society: Nomenclature for the description of sequences variation, <http://www.hgvs.org/mutnomen/>).

Multiple protein alignments were performed using the software SIFT (<http://blocks.fhcrc.org/sift/SIFT.html>) and the search for polymorphisms was carried out using the dbSNP database (<http://www.ncbi.nlm.nih.gov/SNP/>). The effect of substitution on the splicing site was studied using the software NetGene2 (<http://www.cbs.dtu.dk/services/NetGene2/>) and GeneSplicer Web Interface ([http://www.tigr.org/tdb/GeneSplicer/gene\\_spl.html](http://www.tigr.org/tdb/GeneSplicer/gene_spl.html)).

The control population consisted of 50 individuals of European descent, 50 individuals of North African descent, 50 individuals from African descent and 50 individuals of Reunion Island descent, all of whom have no report of pulmonary disease.

## RESULTS

We screened a cohort of 121 patients with diffuse lung disease for the mutation c.218T>C (p.Ile73Thr, I73T). *SFTPC* mutation is present in 18 (17 unrelated) patients and seven novel mutations are described.

The I73T mutation was found in 10 patients without neonatal onset, eight presenting with respiratory viral infection. I73T was inherited in 6 cases. In 4 cases, I73T was sporadic and thus absent in both parents. The 111 patients without I73T mutation were screened for all the coding sequence and the exon-intron boundaries of *SFTPC*. Seven novel *SFTPC* variations were identified (Table 1).

**Table 1. Newly discovered *SFPTC* mutations.**

Nucleotide change <sup>a</sup>	Exon/intron	Amino acid change <sup>b</sup>	Protein domain	Sex	Familial (F)/ <i>de novo</i>	Origin	Age of onset	Virus infection at onset	Outcome
c.116T>C	Exon 2	p.Val39Ala (V39A)	Mature SP-C	F	F	African	4 months	Yes	ILD
c.325-1G>A	Intron 3	-	BRICHOS	M	<i>de novo</i>	North African	Birth	Yes	ILD
c.424delC	Exon 4	p.His142fs	BRICHOS	M	F	North African	Birth	Yes	ILD
c.435G>C	Exon 4	p.Gln145His (Q145H)	BRICHOS	F	<i>de novo</i>	Reunion Island	1 month	No	ILD <sup>#</sup>
c.563T>C	Exon 5	p.Leu188Pro (L188P)	BRICHOS	M	<i>de novo</i>	European	2 months	No	ILD
c.566G>A	Exon 5	p.Cys189Tyr (C189Y)	BRICHOS	M*	F	European	Birth	No	ILD
c.566G>A	Exon 5	p.Cys189Tyr (C189Y)	BRICHOS	F*	F	European	Birth	No	ILD
c.581T>C	Exon 5	p.Leu194Pro (L194P)	BRICHOS	M	F	European	2 months	Yes	ILD

<sup>a</sup>Numbering relative to coding DNA reference sequence NM\_003018.3 and starts at translation initiation site (ATG).

<sup>b</sup>Numbering relative to NP\_003009.2.

M/F: Male/Female

ILD: Interstitial Lung Disease

\*Twins

<sup>#</sup>Death at 3 months

Two mutations, C189Y and L194P, were inherited (Figure 1), whereas the mutations c.325-1G>A and c.424delC were sporadic and absent in the patient's parents and probably appeared *de novo*. Samples of familial individuals of index cases with V39A, Q145H and L188P were unavailable. The novel variants segregated with the disease in the familial cases (C189Y and L194P) and have not been previously reported as polymorphisms, suggesting that they are pathogenic. In addition, each new variant was not detected in 100 control alleles of the same origin. The mutation c.325-1G>A located at the last base of intron 3 is presumed to abolish the normal acceptor splice site using the GeneSplicer prediction software, and to result in the skipping of exon 4 and the deletion of 37 amino acids in the carboxy-terminal domain of



proSP-C. The mutation c.424delC in exon 4 is a frameshift mutation that induces a premature stop codon at position 185. The variation c.435G>C resulting in the substitution of glutamine 145 by histidine, involved the last base of exon 4. This substitution is supposed to abolish the donor splicing site of exon 4 following the software of splice site prediction, and to result in the skipping of exon 4. The V39A mutation is located in mature SP-C. The six other novel variations have been detected in the BRICHOS domain (Figure. 2).

Interestingly, the age of onset was significantly lower in patients with new mutations compared with those carrying the I73T mutation (Table 2). No other significant differences between the two groups of patients were observed.

**Table 2. Characteristics of patients with I73T and new *SFTPC* mutations.**

	<b>Total</b>	<b>I73T mutation</b>	<b>New mutations</b>
<b>Patients n</b>	18	10	8
<b>Sex (Male/Female)</b>	10/8	5/5	5/3
<b>Mutation status</b>			
Familial	10	6	4
<i>de novo</i>	8	4	4
<b>Age at onset (mean±sd in months)</b>	4.0±1.2	5.9±1.9	1.7±0.4*
<b>Viral infection at onset</b>	8	4	4
<b>Neonatal respiratory distress</b>	4	0	4
<b>Symptoms</b>			
Cough	18	10	8
Tachypnea/dyspnea	18	10	8
Cyanosis	18	10	8
Failure to thrive	18	10	8

\*p<0.01, Mann-Whitney

## DISCUSSION

The identification of SP-C alterations has led to significant advances in the diagnosis of diffuse lung disease. Our study on a large cohort of patients confirms that the common mutation I73T in the *SFTPC* gene is a frequent cause of diffuse lung disease. In addition, we described

seven new pathogenic *SFTPC* variants associated with respiratory diseases occurring during the first months of life.

The mutation I73T has been first reported in one allele of an infant with severe respiratory insufficiency and a histological pattern of nonspecific interstitial pneumonia (NSIP) and pulmonary alveolar proteinosis (PAP)<sup>7, 9</sup>. Furthermore, in a cohort of 116 children with ILD, this mutation was present in 7 cases and appeared to be common<sup>12</sup>. The mutation was inherited or observed *de novo*. The phenotype associated with I73T was described as eminently variable, ranging from fatal cases to asymptomatic adult carriers<sup>12, 13</sup>. In our cohort of 121 patients with diffuse lung disease, we found 10 unrelated patients with diffuse lung disease harboring the I73T mutation. This result confirms the high frequency of this mutation in diffuse lung disease.

Besides this frequent mutation, only 12 *SFTPC* mutations have been previously described. We report in this study seven new mutations. The splice site mutation c.325-1G>A is thought to affect the acceptor splice site of exon 4, and may result in the skipping of exon 4. Thus, this mutation probably has the same effect than the previously reported mutation c.435+1G>A affecting the donor site of exon 4<sup>4</sup>. The deletion c.424delC induces the translation of an abnormal protein. The missense mutations V39A, Q145H, L188P, C189Y, and L194P are believed to be pathogenic because they have not been observed in control chromosomes of the same origin. Moreover, the amino acids involved are conserved in mammalian SP-C sequences, and the mutation segregates with the pulmonary disease in familial cases (C189Y and L194P). The phenotype associated with C189Y appeared to be variable, severe in the first months of life and asymptomatic in adults. In contrast, L194P is associated with pulmonary fibrosis in adults. The mutation V39A is the first valine of the polyvaline stretch of mature SP-C. The polyvaline

stretch is believed to be a crucial actor of the secondary structure of SP-C. Indeed, studies of purified mature SP-C have shown that the polyvaline stretch is highly  $\alpha$ -helical and capable of spanning phospholipids bilayers in the liquid crystalline phase<sup>14</sup>.

To date, most of the mutations previously identified in the SP-C gene are localized in the propeptide. In our study, with the exception of the V39A located in the mature peptide, the other new mutations are in the propeptide. Based on functional studies that ascertain the role of *SFTPC* mutations in the pathogenesis of diffuse lung disease, two mechanisms have been suggested<sup>2, 10</sup>. First, owing to a heterozygous mutation, abnormal proSP-C could disrupt the regular processing of proSP-C within the cell and consequently inhibit SP-C production in the alveoli. This is the case for the I73T and E66K mutations altering intracellular trafficking<sup>7, 10</sup>. Also, as shown in the  $\Delta$ exon4 and L188Q mutations, aberrant proSP-C forms aggregates that could be toxic for alveolar type II cells and induce an unresolved inflammatory response<sup>15-17</sup>. Moreover, in two patients, diffuse lung disease have also been associated with SP-C deficiency without any detectable *SFTPC* mutation<sup>9</sup>. Thus, Alteration of SP-C metabolism or complete SP-C deficiency is critical in diffuse lung disease pathophysiology.

Interestingly, we found three novel variants, L188P, C189Y and L194P, located in the C-terminus of proSP-C BRICHOS domain (F94-I197)—a highly conserved domain. The BRICHOS domain is composed of about 100 amino acids in the proprotein. It is found in several proteins associated with degenerative and proliferative diseases<sup>18</sup>. The BRICHOS domain is thought to be involved in protein processing and is assumed to have a chaperone-like function. It plays a critical role in proSP-C targeting and processing<sup>16, 19, 20</sup>. Indeed, in previous studies, two patients have been found to carry a mutation encoding substitution (L188R and L188Q) of the same conserved leucine residue altered in our cohort (L188P). The L188R

mutation has been detected in one infant with ILD <sup>5</sup>. On the other hand, the L188Q mutation is detected in large familial pulmonary fibrosis kindred, either in children or adults with cellular interstitial pneumonitis or usual interstitial pneumonitis <sup>6</sup>. We have observed L188P in an 18-year-old patient who had diffuse lung disease during the first year of life. Unfortunately, familial samples were not available.

Descriptions of the reported new mutations will be the subject of further functional studies, which will improve knowledge on the pathogenesis of associated surfactant diffuse lung disease in children.

#### **ACKNOWLEDGEMENTS**

We thank the patients and their families for their cooperation in this study. We also thank Catherine Meunier, France Michel, Corinne Chauve, Isabelle Sargis, Magali Niasme, and Cristina Das Neves for their expert technical assistance. This work was supported by the « Département de la Recherche Clinique et du Développement » (PHRC 2007, « “Surfactant disorders associated with chronic lung disease in children” »).

"The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in Journal of Medical Genetics and any other BMJ PGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence (<http://JMG.bmj.com/misc/ifora/licenceform.shtml>)."

## REFERENCES

1. Deutsch GH, Young LR, Deterding RR, *et al.* Diffuse lung disease in young children: application of a novel classification scheme. *Am J Respir Crit Care Med* 2007;**176**(11):1120-8.
2. Hartl D, Griese M. Interstitial lung disease in children -- genetic background and associated phenotypes. *Respir Res* 2005;**6**:32.
3. Nogee LM. Alterations in SP-B and SP-C expression in neonatal lung disease. *Annu Rev Physiol* 2004;**66**:601-23.
4. Nogee LM, Dunbar AE, 3rd, Wert SE, *et al.* A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *N Engl J Med* 2001;**344**(8):573-9.
5. Nogee LM, Dunbar AE, 3rd, Wert S, *et al.* Mutations in the surfactant protein C gene associated with interstitial lung disease. *Chest* 2002;**121**(3 Suppl):20S-21S.
6. Thomas AQ, Lane K, Phillips J, 3rd, *et al.* Heterozygosity for a surfactant protein C gene mutation associated with usual interstitial pneumonitis and cellular nonspecific interstitial pneumonitis in one kindred. *Am J Respir Crit Care Med* 2002;**165**(9):1322-8.
7. Brasch F, Griese M, Tredano M, *et al.* Interstitial lung disease in a baby with a de novo mutation in the SFTPC gene. *Eur Respir J* 2004;**24**(1):30-9.
8. Hamvas A, Nogee LM, White FV, *et al.* Progressive lung disease and surfactant dysfunction with a deletion in surfactant protein C gene. *Am J Respir Cell Mol Biol* 2004;**30**(6):771-6.
9. Tredano M, Griese M, Brasch F, *et al.* Mutation of SFTPC in infantile pulmonary alveolar proteinosis with or without fibrosing lung disease. *Am J Med Genet A* 2004;**126A**(1):18-26.
10. Stevens PA, Pettenazzo A, Brasch F, *et al.* Nonspecific interstitial pneumonia, alveolar proteinosis, and abnormal proprotein trafficking resulting from a spontaneous mutation in the surfactant protein C gene. *Pediatr Res* 2005;**57**(1):89-98.
11. Soraisham AS, Tierney AJ, Amin HJ. Neonatal respiratory failure associated with mutation in the surfactant protein C gene. *J Perinatol* 2006;**26**(1):67-70.
12. Cameron HS, Somaschini M, Carrera P, *et al.* A common mutation in the surfactant protein C gene associated with lung disease. *J Pediatr* 2005;**146**(3):370-5.
13. Percopo S, Cameron HS, Nogee LM, *et al.* Variable phenotype associated with SP-C gene mutations: fatal case with the I73T mutation. *Eur Respir J* 2004;**24**(6):1072-3.
14. Beers MF, Mulugeta S. Surfactant protein C biosynthesis and its emerging role in conformational lung disease. *Annu Rev Physiol* 2005;**67**:663-96.
15. Bridges JP, Wert SE, Nogee LM, *et al.* Expression of a human surfactant protein C mutation associated with interstitial lung disease disrupts lung development in transgenic mice. *J Biol Chem* 2003;**278**(52):52739-46.
16. Mulugeta S, Nguyen V, Russo SJ, *et al.* A surfactant protein C precursor protein BRICHOS domain mutation causes endoplasmic reticulum stress, proteasome dysfunction, and caspase 3 activation. *Am J Respir Cell Mol Biol* 2005;**32**(6):521-30.
17. Wang WJ, Mulugeta S, Russo SJ, *et al.* Deletion of exon 4 from human surfactant protein C results in aggresome formation and generation of a dominant negative. *J Cell Sci* 2003;**116**(Pt 4):683-92.
18. Sanchez-Pulido L, Devos D, Valencia A. BRICHOS: a conserved domain in proteins associated with dementia, respiratory distress and cancer. *Trends Biochem Sci* 2002;**27**(7):329-32.
19. Mulugeta S, Maguire JA, Newitt JL, *et al.* Misfolded BRICHOS SP-C mutant proteins induce apoptosis via caspase-4- and cytochrome c-related mechanisms. *Am J Physiol Lung Cell Mol Physiol* 2007;**293**(3):L720-9.
20. Nerelius C, Martin E, Peng S, *et al.* Mutations linked to interstitial lung disease can abrogate anti-amyloid function of prosurfactant protein C. *Biochem J* 2008.

## TABLE AND FIGURES LEGENDS

**Table 1:** Newly discovered *SFPTC* mutations.

**Table 2:** Characteristics of patients with I73T and new *SFPTC* mutations.

**Figure 1:** Pedigree of families with the new *SFPTC* mutations C189Y and L194P.

**Figure 2:** Localization of new *SFPTC* Mutations. Schematic representation of *SFPTC* gene.

(a) (exon I to VI), corresponding mRNA (b), pro-SP-C (c) and SP-C (d) protein sequence with annotated newly identified mutations.

Figure. 1

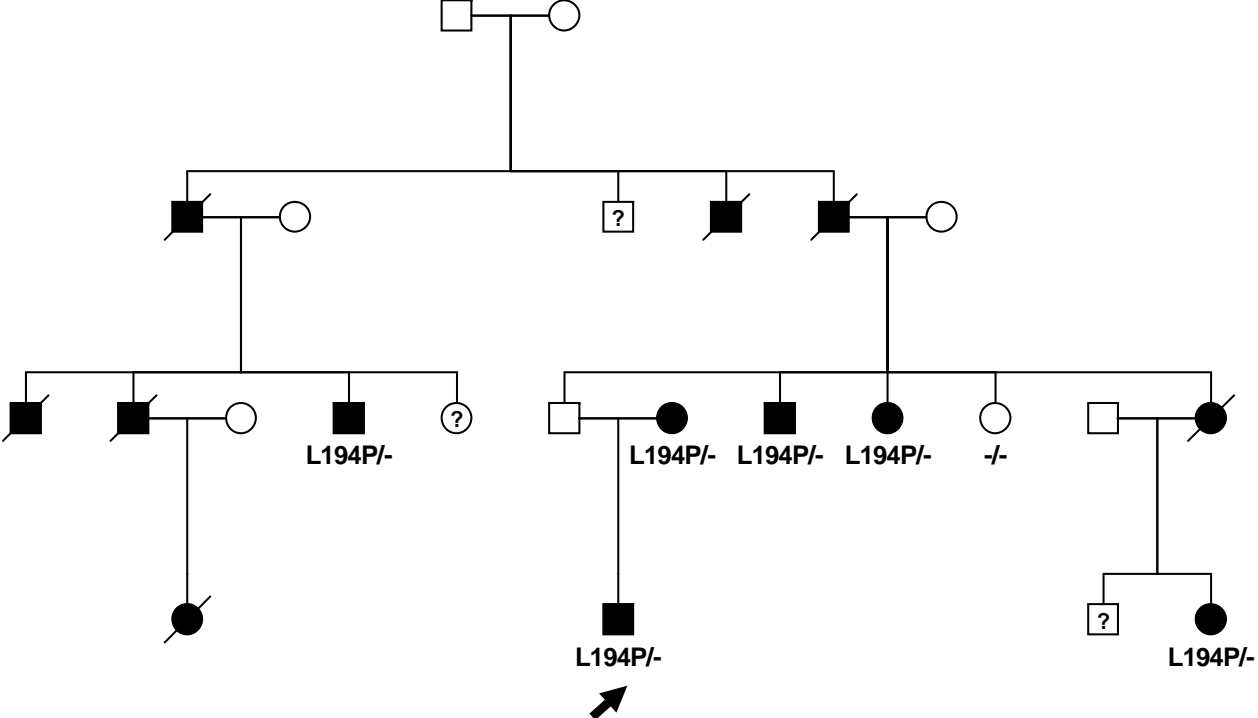
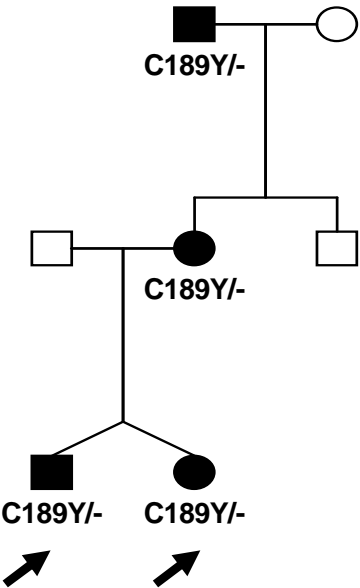


Figure. 2

