EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original article

An atypical pulmonary fibrosis is associated with co-inheritance of mutations in the calcium binding protein genes *S100A3* and *S100A13*

E.A. Al-Mutairy, F. Imtiaz, M. Khalid, S. Al Qattan, S. Saleh, L. Mahmoud, M. Al-Saif, L. Al-Haj, A. Al-Enazi, A.M. AlJebreen, S. Mohammed, A. Mobeireek, K. Alkattan, M.A. Chisti, I.G. Luzina, M. Al-Owain, I. Weheba, A. Abdelsayed, K. Ramzan, J.L. Janssen, W. Conca, A. Alaiya, K.S. Collison, B.F. Meyer, S.P. Atamas, K.S. Khabar, J.D. Hasday, F. Al-Mohanna

Please cite this article as: Al-Mutairy EA, Imtiaz F, Khalid M, *et al*. An atypical pulmonary fibrosis is associated with co-inheritance of mutations in the calcium binding protein genes *S100A3* and *S100A13*. *Eur Respir J* 2019; in press (https://doi.org/10.1183/13993003.02041-2018).

An atypical pulmonary fibrosis is associated with co-inheritance of mutations in the calcium binding protein genes *S100A3* and *S100A13*

Al-Mutairy EA[1,2,3], Imtiaz F[4], Khalid M[1], Al Qattan S[2], Saleh S[2], Mahmoud L[5], Al-Saif M[5], Al-Haj L[5], Al-Enazi A[2], AlJebreen AM[6], Mohammed S[7], Mobeireek A[1], Alkattan K[1,3], Chisti MA[8], Luzina IG[9,10], Al-Owain M[3,11], Weheba I[1,12], Abdelsayed A[1,13], Ramzan K[4], Janssen J L[14], Conca W[1,2,3], Alaiya A[15], Collison KS[2], Meyer BF[4], Atamas SP[9,10], Khabar KS[3,5], Hasday JD[9,10], Al-Mohanna F[2,3]

[1]Department of Medicine, King Faisal Specialist Hospital and Research Centre (KFSHRC), Riyadh, 11211, Saudi Arabia.[2]Department of Cell Biology, KFSHRC, Riyadh, 11211, Saudi Arabia.[3]College of Medicine, Al-Faisal University, Riyadh, Saudi Arabia.[4]Department of Genetics, KFSHRC, Riyadh, 11211, Saudi Arabia.[5]BioMolecular Medicine, KFSHRC, Riyadh, 11211, Saudi Arabia.[6]Department of Radiology, KFSHRC, Riyadh, 11211, Saudi Arabia. [7]Department of Pathology and Laboratory Medicine, KFSHRC, Riyadh, 11211, Saudi Arabia.[8]Department of Dermatology, KFSHRC, Riyadh, 11211, Saudi Arabia.[9]University of Maryland School of Medicine, Baltimore, MD 21201, U.S.A. [10]Baltimore VA Medical Center, Baltimore, MD 21201, U.S.A. [11]Department of Medical Genetics, KFSHRC, Riyadh, 11211, Saudi Arabia.[12]National Research Centre, Cairo, Egypt.[13]Ain Shams University, Cairo, Egypt. [14]St. Joseph's Hospital& Department of Medicine, McMaster University, Ontario, Canada, [15]Stem Cell Therapy Program, KFSHRC, Riyadh, 11211, Saudi Arabia.

Correspondence

Eid Al-Mutairy, MD
Departments of Medicine and Cell Biology
King Faisal Specialist Hospital and Research Centre
Riyadh 11211
Saudi Arabia
almutairyeid@gmail.com

Futwan Al-Mohanna, PhD
Department of Cell Biology
King Faisal Specialist Hospital and Research Centre
Riyadh 11211
Saudi Arabia
futwan@kfshrc.edu.sa

Abstract:

Background: Pulmonary fibrosis is one of the leading indications for lung transplantation. The disease, which is of unknown etiology, can be progressive resulting in distortion of extracellular matrix (ECM), inflammation, fibrosis and eventual death.

Patients and Methods: Thirteen patients born to consanguineous parents from two unrelated families presenting with interstitial lung disease were clinically investigated. Nine patients developed respiratory failure and subsequently died. Molecular genetic investigations were performed on patients' whole blood or archived tissues and cell biological investigations were performed on patient-derived fibroblasts.

Results: The combination of a unique pattern of early onset lung fibrosis (12-15 years old), distinctive radiological findings, including (1) traction bronchiectasis (2) intralobular septal thickening (3) shrinkage of the secondary pulmonary lobules mainly around the bronchovascular bundles and (4) early type 2 respiratory failure (elevated blood CO_2 levels) represents a novel clinical subtype of familial pulmonary fibrosis. Molecular genetic investigation of families revealed a hypomorphic variant in *S100A3* and a novel truncating mutation in *S100A13*, both segregating with the disease in an autosomal recessive manner. Family members that were either heterozygous carriers or wild-type normal for both variants are unaffected. Analysis of patient-derived fibroblasts demonstrated significantly reduced S100A3 and S100A13 expression. Further analysis demonstrated aberrant intracellular calcium homeostasis, mitochondrial dysregulation and differential expression of ECM components.

Conclusion: Our data demonstrate that digenic inheritance of mutations in *S100A3* and *S100A13* underlie the pathophysiology of pulmonary fibrosis associated with a significant reduction of both proteins, which suggests a calcium dependent therapeutic approach for management of the disease.

Introduction

Interstitial lung diseases (ILDs) are a heterogeneous group of disorders of largely unknown etiology that are characterized by variable types of interstitial and alveolar inflammation, parenchymal remodeling, and fibrosis [1]. The most common ILD is idiopathic pulmonary fibrosis (IPF), a progressive disorder that usually affects individuals over 55 years of age. Because of the lack of effective treatments and the rapid progression to respiratory failure and death, pulmonary fibrosis (PF) remains one of the leading indications for lung transplantations worldwide [2, 3]. Several environmental risk factors have been implicated in the pathogenesis of IPF. A genetic predisposition has been demonstrated [4], however the great majority of cases of IPF are sporadic.

Familial pulmonary fibrosis (FPF), defined as idiopathic interstitial lung disease in two or more first-degree relatives (parent, sibling, or offspring), has been attributed to non-synonymous monogenic mutations in surfactant protein A2 (*SFTPA2*), surfactant protein C (*SFTPC*), or ATP-binding cassette A3 (*ABCA3*) [5, 6]. In addition, a common variant (rs35705950) in the promoter of the gene encoding mucin 5B (*MUC5B*) substantially increases its expression and markedly increases the risk of PF [7]. These mutations are proposed to converge on activation of the unfolded protein response [8]. Amongst a plurality of FPF kindred, ~15%, had mutations in telomerase genes *TERT* and *TERC,* and exhibited shortened telomeres [9, 10]. Telomere shortening was also evident in 25% of patients with sporadic IPF who did not have identifiable mutations in *TERT* or *TERC* [11]. Several rare variants of other genes related to telomere maintenance, such as *TINF2*, *DKC1*, *RTEL1*, *PARN* and *NAF1*, have also been identified in ~25% of patients with FPF [12].

Despite these advances, the pathogenesis of sporadic PF remains unclear. Here we have identified two variants in the calcium binding protein genes *S100A3* (NM_002960) and *S100A13* (NM_001024210) segregating with the disease in seven siblings screened from two unrelated families with PF. The mutations led to lower expression of the proteins, aberrant receptor-mediated intracellular calcium responses, reduced capacity to tolerate external oxidative stress and altered extracellular matrix (ECM) protein expression in cells isolated from patients. These results will improve our understanding of the pathogenesis of fibrosing lung diseases.

Materials and Methods

All samples were collected with Institutional Review Board approved written informed consents. The study was approved by the Research Advisory Council (RAC), King Faisal Specialist Hospital and Research Centre (KFSH&RC-RAC# 2120 009).

Sequence analysis

Genomic DNA was extracted from whole blood or paraffin-embedded archived tissue of the affected patients, their parents and unaffected relatives (Families 1 & 2) using standard methods. Bi-directional sequencing of the coding regions of known IPF-associated genes (*TERT*, *TERC*, *ABCA3* and *SFTPB*) was performed in patients and nuclear family members of Family 1A. The full coding regions of *S100A3* (NM_002960.1) and *S100A13* (NM_001024210) were sequenced for all available members (affected and unaffected of Families 1 &2) and 28 patients with sporadic PF, using standard PCR conditions. Sequence analysis was performed manually using the SeqMan 6.1 module of the Lasergene software package (DNA Star Inc. WI, USA).

Linkage analysis and homozygosity mapping

Genotyping of all available family members using the Affymetrix Axiom® Genome-Wide CEU 1 Array platform was performed and analyzed for homozygosity mapping using AutoSNPa. Linkage analysis was performed using the Allegro module of the easyLINKAGE [13]. Direct sequencing of candidate genes in the linkage interval and exome re-sequencing data of genomic DNA was performed using primer pairs designed to cover the entire coding region of each gene.

Whole Exome Sequencing

Whole exome sequencing (WES) was first performed on the affected son of Family 1A using the Illumina® HiSeq2000 platform with TruSeqv3 chemistry by preparing and enriching the sample according to the manufacturer's standard protocol instructions. Concentration of each library was determined using Agilent's (Agilent Technologies, Santa Clara, CA, USA) QPCR NGS Library Quantification Kit (G4880A) and the sample was sequenced at a final concentration of 10 nM. Mapping and alignment was performed on read files (Fastq) generated from the sequencing platform via the manufacturer's proprietary software and using human genome (hg19/b37) and the Burrows-Wheeler Aligner (BWA) package, version 0.6.1 [14]. Further realignment and variant analysis were performed eventually determining SNP novelty against dbSNP (Human Build 135) [15-17]. Variants were annotated with gene and gene function from Ensembl (http://www.ensembl.org/index.html) [18] and further analysis of possible causative variants by filtering the full exome dataset for all deletions, insertions, nonsense and canonical splice-site mutations, as well as missense mutations (with a PhyloP score of >3.5 of the underlying base change) were determined and reported. Similar WES analysis was performed later on one affected patient from both families 1B & 2 using the same methodology.

Tissue Culture

Skin biopsies were collected from two patients and two controls according to IRB approved protocols. Fibroblasts were isolated and cultured as described previously [19].

Immunofluorescence and Western blotting

Immunofluorescence staining was performed on lung tissue sections prepared from paraffin blocks collected according to IRB approved protocols from probands and healthy control lungs (healthy donor lungs before transplantation). Primary antibodies to S100A3 and S100A13 (abcam, Abdulla Fouad Company, Saudi Arabia) and fluorescein-conjugated secondary anti-rabbit IgG (ThermoFisher Scientific, USA) were used according to manufacturer's instruction. For fibroblast immunofluorescence, anti-S100A3 was purchased from Santa Cruz (Santa Cruze, CA). Images were acquired and analysed using cellSens Dimension software 1.9. (Olympus Corporation GmbH, Germany). Image acquisition was thresholded against controls where the primary antibody was omitted and fluorescence intensity due to secondary antibody alone was considered non-specific.

For Western blots, cells were lysed, separated on either 7.5% SDS-PAGE (Bio-Rad, Hercules, CA) or 4-12% gradient SDS-PAGE and transferred onto PVDF membranes (Life Technologies, Carlsbad, CA) or nitrocellulose membranes (Amersham Hybond ECL nitrocellulose membranes), and immunoblotted using primary rabbit antibodies against *S100A3* (Santa Cruze, CA), *S100A13*, MMP2, MMP9, TIMP1, actin (Santa Cruze, CA) or GAPDH (Cell Signaling; Danvers, MA); followed by peroxidase-conjugated goat anti-rabbit IgG (Jackson ImmunoResearch Laboratories; West Grove, PA). Bands were visualized using chemiluminescence (SuperSignal West Pico, Thermo Scientific/Pierce, Rockford, IL) and exposed to X-ray film or scanned using ChemiDoc™ XRS+ Molecular Imager with Image lab™ software (Bio-Rad Laboratories, Inc. USA). Western blot membranes were scanned and bands were processed using Count and Measure function (cellSens Dimension software 1.9) and intensities of appropriate regions of interest (ROI) were determined and expressed relative to corresponding β-actin band intensities.

Sample preparation for label-free protein in-solution digestion: Cells derived from patient and control samples were lysed using the RapiGest MS compatible lysis buffer (Waters, Manchester, UK). A total of 100 µg protein from whole cell lysate of each sample was subjected to proteome analysis by in-solution tryptic digestion as previously described [20, 21].

Protein identification, by LC-MS^E SynaptG2 Platform

We used label-free quantitative 1-dimensional Nano Acquity liquid chromatography tandem mass spectrometry on Synapt G2 (Waters, Manchester, UK) and generated expression protein profiles between the sample groups. The instrument settings were optimized as described previously [20, 22, 23].

Expression proteomics data analysis and bioinformatics

All data acquisitions were in triplicate runs with automated data processing and database search using the Uniprot Human specific protein sequence database on the Progenesis QI for proteomics for protein identification platform (Waters, UK and Nonlinear Dynamics, Newcastle, UK). Normalized protein abundance of significantly regulated proteins were considered (ANOVA; $p < 0.05$), and significant expression was defined as a fold change >1.5 and a false discovery rate (FDR) ~3%. The significant expression dataset was further evaluated for their functional/signaling pathway implications using the Ingenuity Pathway Analysis v8.7 (IPA Ingenuity Systems, www.ingenuity.com).

Quantitative reverse transcriptase-polymerase chain reaction (qRT-PCR)

Total cellular RNA was isolated using TRIzol reagent (Ambion, Grand Island, NY), and complementary DNA was synthesized from 1–5 μg of RNA using a RT^2 First Strand cDNA kit (Qiagen, Germantown, MD) according to the manufacturer's protocol. Primers for 18S rRNA were purchased from SABiosciences/Qiagen (Valencia, CA). Primers for *S100A3* (forward: 5'-CCCGAACTGGTCAACTCTCA; reverse: 5'- GCCTGGCAGAGCTTGTATTT), and plasmid backbone (forward: 5'-GTGGCGCTTTCTCATAGCTC; forward: 5'-TGTCTTACCGGGTTGGACTC) and for *S100A13* (forward: 5'-CATCTGCTCAAGGATGTGGG; reverse: 5'-TCCTGATCTTCAGGTCTTT) were designed using PRIMER3. SABiosciences/Qiagen (Valencia, CA). qRT-PCR was performed on an Applied Biosystems StepOne Plus PCR system (Carlsbad, CA) using RT2 SYBR Green qRT-PCR Mastermix (SABiosciences) according to the manufacturer's directions. For qRT-PCR of *S100A3* in control and patient samples the following primers were used: (forward: 5'-GGACCCCGACTGAGTTTCG; reverse: 5'-GCTCTGAGGGGCAGTCCTTG) and for *S100A13* (forward: 5'-CATCTGCTCAAGGATGTGGG-3'; reverse: 5'-TCCTGATCTTCAGGTCTTT-3'. For GAPDH the following primers were used (forward: 5'- CACCATCTTCCAGGAGTGAG; reverse: 5'-TCACGCCACAGTTTCCCGGA).

Intracellular calcium and mitochondrial integrity measurements

Cytosolic calcium measurements were performed on patient or control fibroblasts (from unaffected donors) as described previously [24]. Mitochondrial calcium measurements were performed using Rhod-2 AM as described previously [19]. Receptor mediated changes in intracellular fluorescence intensity in response to fibroblast growth factor- (FGF)-2 (Sigma USA, 10ng/ml), bradykinin (Sigma USA, 50μM) and ionomycin (Sigma USA, 2μM) were followed using Zeiss LSM 510 META laser scanning confocal

system (Carl Zeiss MicroImaging, GmbH, Germany). Mitochondrial staining was performed using Mito Tracker® Red CMXRos (invitrogen™ Molecular Probes™, USA) (1µM, 5 min at 37°C) and viewed under Zeiss Yokogawa Spinning Disk confocal microscopy system (Carl Zeiss MicroImaging, GmbH, Germany).

Transmission Electron Microscopy (TEM) and Flow Cytometry

For TEM, cells were fixed with 2.5% glutaraldehyde in cacodylate buffer (0.1 M, pH 7.4) for a minimum of 48 h. Osmication was performed using reduced osmium (1:1 mixture of 2% osmium tetroxide and 3% potassium ferrocyanide). After pre-embedding in 1% agar, samples were dehydrated in ethanol series and embedded in epoxy resin. Thin sections (70 to 100 nm thickness) were collected on copper grids and contrasted with lead citrate. Imaging was performed using a transmission electron microscope operating at 300 kV (Titan Cryo Twin, FEI Company, Hillsboro, OR). Images were recorded on a 4k x 4k CCD camera (Gatan Inc., Pleasanton, CA). For flow cytometry, cells (1×10^6 cells/ml) were labeled with Mito Tracker® Green FM (1µM) for 45 min on ice, washed (PBS, pH 7.2), fixed in 1% paraformaldehyde, and analyzed using a BD FACSCalibur flow cytometer (BD Biosciences).

Statistical Analysis

One-way analysis of variance (ANOVA) with Holm-Sidak's multiple comparison test was used to measure statistical significance using GraphPad prism software (GraphPad Software, Inc. USA). Unpaired two-tailed Student *t-test* was used when appropriate. A *p value of* ≤ 0.05 was considered significant.

Results

Brief Case Description

Family 1A: Three siblings (Figure 1A), one boy (F1: IV-1) and two girls (F1: IV-2 & IV-3) from a total of seven children were born healthy to consanguineous parents (F1:III-6 & 7) after normal pregnancies and deliveries. The parents did not report any developmental delay or medical problems until all three children developed PF at a young age. The boy developed PF at 12 years of age and both girls at 13 years of age. All three affected siblings had an identical clinical presentation and course (Figure 1B). All developed dyspnea in their early teens and, aside from the lung abnormalities, medical examinations did not reveal any abnormalities in appearance, developmental milestones, laboratory findings, or other organ systems. Fibrosis was revealed by chest CT imaging and characterized by traction bronchiectasis, intralobular septal thickening, and shrinkage of the secondary pulmonary lobules mainly around the bronchovascular bundles. The core of the lungs was extensively involved suggesting a bronchocentric interstitial fibrosis (Figures 1B & 1C) and pulmonary function testing indicated severe restriction and impaired oxygen transfer (Figure 1F). Extensive medical investigations ruled out autoimmune and occupational causes of PF. Respiratory symptoms worsened with age, and all 3 patients developed chronic type 2 respiratory failure (elevated blood CO_2 levels [25]), requiring non-invasive ventilation. Two of the siblings underwent lung transplantation at ages 23 years (F1: IV-1) and 22 years (F1: IV-2), but died due to primary graft dysfunction (PGD). Both received routine immunosuppression and their cross match was negative. The third sibling (F1: IV-3) died from respiratory failure at 23 years of age without having a lung transplant. Analysis of lung tissue obtained at autopsy from one of the subjects (F1: IV-2) revealed areas of interstitial inflammation, mild fibrosis involving the alveolar walls, mild subpleural fibrosis, and advanced fibrosis sparing of the subpleural space (Figure 1G**).**

Family 1B: Four affected girls from a total of nine children born healthy to consanguineous parents (F1: III-8 & 9) were enrolled. The mother of this family is the second cousin of both of the parents of Family 1A. Two of the girls (Figure 1A; F1: IV-9 and IV-10) died of respiratory failure in their late twenties. Detailed clinical information is not available as they died prior to the family presenting to our clinic. Similar to Family 1A, an extensive workup of the living patients currently aged 31 (F1: IV-12) and 34 years (F1: IV-15), with symptoms beginning at age 12 and 15, respectively, revealed no known autoimmune and occupational causes of PF. CT imaging revealed identical findings of fibrosis and pulmonary function tests showed severe restriction pattern (images not shown). No developmental or extra-pulmonary findings were documented. Lung transplantation work up has been initiated for both sisters.

Family 2: A third family presenting with a strikingly similar clinical pattern to Family 1 (A & B). Family 2 comprised of a total of 10 children with six affected siblings (4 girls and 2 boys) born to healthy consanguineous parents. Extensive pedigree analysis confirmed that they were unrelated to Family 1. Four of the affected children (2 girls and 2 boys) had died previously due to lung disease (between the ages of 25 and 32). Detailed clinical information for the deceased is not available. Both of the surviving sisters developed respiratory symptoms at 13 years of age in one (F2: IV-7) and estimated as late teenage years in the second (F2: IV-8). For both sisters, pulmonary function tests showed a severe restriction pattern. CT imaging was very similar to Family 1 (Figure 1D-F & H). Like the affected individuals in Family 1, the affected sisters in Family 2 showed normal appearance, development and laboratory findings, and autoimmune and environmental causes of lung fibrosis were excluded. One of the sisters underwent a lung transplant at age 27 (F2: IV-7) and is currently 35 years and in good health. The second sister (F2: IV-8) is currently on the waiting list for lung transplantation.

Identification of a novel PF region on chromosome 1

Homozygozity mapping using families 1 and 2 defined a single 34 Mb (chr1:120,127,864-154,749,047 bp) critical interval shared by all 7 affected individuals (F1: IV-1, 2 & 3, F1:IV-12 & 15 and F2: IV:7 &8) (Figure 2A), defined proximally by rs10802117 and distally by rs11808053 containing over 363 annotated genes (Supplementary Figure 1B). A combined parametric multipoint linkage analysis using a total of 17 individuals (7 affected and 10 unaffected) from F1 and F2 revealed a single significant peak with a maximum LOD score of 5.28 corresponding to chromosome 1p12-q21.3 (Figure 2B). Initial Sanger sequencing of 3 candidate genes in this region (based upon function and/or association with ILD) failed to identify a causative mutation.

Exome sequencing

Initially, whole exome sequencing was performed in the proband of Family 1A (F1: IV-1, Figure 1A). After filtering for homozygous non-synonymous SNVs within the linkage interval, that were either novel or had either low or unknown minor allele frequency in dbSNP, only 3 previously described variants, rs3795737 in *ISG20L2*, rs143224912 in *SETDB1*, and rs138355706 in *S100A3* and one novel variant in *S100A13*, were identified. The *ISG20L2* and *SETDB1* variants were excluded based on their frequencies in normal population cohorts. Sanger sequencing of Family 1 showed that both rs138355706 in *S100A3* (229C>T, missense causing a p.R77C mutation) and a four base-pair deletion in *S100A13* (c.238-241delATTG causing a frameshift p.I80Gfs*13) segregated completely with ILD in family 1 based upon recessive inheritance (Figure. 2C), were in total linkage disequilibrium, and were present in a *cis* conformation. Allele frequency for rs138355706 within the population was calculated from exome sequencing of 2000 individuals and was found to be 0.1% (Saudi Genome Project, unpublished data). Furthermore 500 ethnically-matched normal controls were genotyped for this change by re-sequencing of

S100A3. Three of these individuals were heterozygous, but none was homozygous for rs138355706. Sequencing of the *S100A3* intronic and 5' flanking sequences was performed in the affected patients and no other variants were identified (data not shown). The novel truncated variant in *S100A13* was not found in our "in-house" Saudi exome data (unpublished from the Saudi Human Genome Project), 1000 genome and gnomAD databases. The c.229C>T (p.R77C) variant in *S100A3* and c.238-241delATTG (p.I80Gfs*13) mutation in *S100A13* also segregated fully with ILD in Families 1B and 2.

Haplotype analysis

Haplotype analysis carried out using 8 markers (4 microsatellite markers flanking *S100A3, S100A13* and 3 further intragenic markers) (Supplementary Figure 1A) confirmed that all affected individuals from both families shared a specific disease haplotype on both chromosomes that was not present in the unaffected individuals, suggesting a shared extended haplotype from a common founder.

Consequences of the *S100A3* c.229 C>T and *S100A13* variants

The *S100A3* c.229 C>T variant resulted in an arginine to cysteine substitution at residue 77 within the second of the two EF-hand calcium-binding motifs in the protein. The predicted impact of this variant on protein structure/function was evaluated using the Polyphen-2 (ver. 2.2.2), SIFT prediction, and Combined Annotation Dependent Depletion (CADD) programs. Polyphen-2 and SIFT programs predicted minor effects of the mutation on protein structure/function with scores of 0.004 and 0.21, respectively, and a CADD PHRED score of 17.65. Since microRNAs can modify translation efficiency by binding to coding sequence as well as 3'UTR sequence, we sought to determine whether the c.229 C>T mutation altered any known microRNA binding sites using MicroSNiPer (release 19) and a minimum 7-nucleotide seed sequence. We found no effect of the SNV on predicted microRNA binding sites. Moreover, analysis of predicted protease cleavage sites using PeptideCutter™ did not reveal any effect on protease cleavage sites in *S100A3*. However, the Human Splicing Finder program (Version 3.1, http://www.umd.be/HSF3/) predicted an alteration in the exonic splicing enhancer (ESE) of *S100A3* [26] and creation of a new exonic splicing silencer (ESS) site, which is known to play a role in constitutive and alternative splicing [27, 28]. The *S100A13* loss of function variant is, as expected, predicted to be "disease causing" by MutationTaster.

Effect of mutations on *S100A3* and *S100A13* expression

Indirect immunofluorescence staining of thin sections obtained from healthy lung tissues (donor lungs before transplantation) demonstrated the presence of both proteins (Figure 3A&B). The staining was consistent with reports in *The Human Protein Atlas* demonstrating the expression of S100A3 and S100A13 in the lungs. (https://www.proteinatlas.org/ENSG00000188015-S100A3/tissue and https://www.proteinatlas.org/ENSG00000189171-S100A13/tissue). In normal lung tissue, S100A3 and

S100A13 expression was distributed along the ciliary edges and the apical surfaces of the bronchiolar epithelium of the small airways. In contrast with control lung tissue and lung tissue from a patient with sporadic IPF, the lung tissue from affected family members demonstrated greatly reduced S100A3 and S100A13 expression. (Figure 3A&B). Low S100A3 and S100A13 expression levels were also found in patient-derived skin fibroblasts when compared with skin fibroblasts from normal controls (Figure 3C, upper panel). Western blotting confirmed significantly reduced expression of S100A3 and S100A13 in skin fibroblasts isolated from patient cells compared with cells from healthy controls (Figure 3C; lower panel). The reduced expression of the mutant proteins in patients was paralleled by reduced expression of mRNA levels of both variants compared to control (Figure 3D).

Effect of *S100A3* and *S100A13* mutations on intracellular calcium signaling and mitochondrial structure and function

Since the *S100A3* and *S100A13* genes encode calcium-binding proteins, we explored the possibility that the mutations may affect intracellular calcium homeostasis. We measured intracellular calcium changes $[Ca^{++}]_i$ in response to bradykinin (50 μM) and fibroblast growth factor (FGF)-2 (10 ng/ml) in skin fibroblasts isolated from healthy controls and patients. Receptor-mediated calcium release was significantly reduced in fibroblasts from patients compared with control fibroblasts (Figure 4A&C). The increase in calcium signal following bradykinin stimulation was reduced in patient cells compared with controls (1.47±0.14-fold vs. 2.34 ± 0.07-fold; $p<0.0001$) (Figure. 4B). FGF-2 induced a 1.19 ± 0.02-fold increase in calcium signal in patient cells vs 1.41 ± 0.06-fold increase in control cells ($p=0.017$) (Figure. 4D). Ionomycin-induced calcium release was also significantly reduced in patient fibroblasts compared with control fibroblasts ($p<0.02$) (Figure. 4E). Since mitochondria accumulate calcium and shape the temporal and spatial calcium changes in many cell types [29, 30], we analyzed the bradykinin-induced increase in intramitochondrial calcium levels, which was reduced in patient fibroblasts compared with normal control fibroblasts (Figure. 4F, $p=0.015$). Furthermore, compared with mitochondria in control cells, the mitochondria in patient cells exhibited aberrant morphology and a more punctate Mito Tracker® Red CMXRos fluorescence pattern (Figure. 5A). In addition, patient cells appear to have increased mitochondrial staining than control cells (Figure. 5A), which was confirmed by flow cytometery analysis of Mito Tracker Green-stained cells ($p=0.002$) (Figure. 5B). Transmission electron microscopy (TEM) showed evidence of mitochondrial damage with reduced cristae and reduced rough endoplasmic reticulum (RER) in patient samples compared with control (Figure. 5C). The functional integrity of the mitochondria in both patient and control cells was further investigated by measuring the effect of external oxidative stress induced by treating the cells with hydrogen peroxide (0.03%). Figure 5D illustrates the ability of control cells to resist oxidative stress compared to patient cells.

Effect of *S100A3 and S100A13* mutation on extracellular matrix remodeling

Since pulmonary extracellular matrix (ECM) remodeling is the hallmark of IPF [31, 32], we compared expression of matrix-related proteins in patient and control fibroblasts. Western blot analysis demonstrated increased expression of two matrix metalloproteinases, MMP2 and MMP9, and reduced expression of tissue inhibitor of matrix metalloproteinase (TIMP)-1 in patient fibroblasts compared with control cells (Figure. 6A). Proteomic analysis of ECM components demonstrated increased expression of MMP1,3 and 14 in patient fibroblasts compared with control cells (Figure. 6B). This was paralleled by differential expression of collagens I (COL1A2), collagen VI (COL6A1&2) collagen VIII (COL8A1), collagen triple helix repeat-containing protein 1 (CTHRC1) and procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 (PLOD1) (Figure. 6C).

Discussion

S100A3 and S100A13 are members of a large family of acidic, low molecular weight calcium binding proteins. More than 20 different members have been identified with diverse functions [33-36]. The proteins are found exclusively in vertebrates. S100 family members exist as homodimers, heterodimers and multi-oligomers, displaying cell- and tissue-specific patterns of expression [37] with significant structural similarities to calmodulins [38, 39]. Although the proteins function intracellularly, secreted members exhibit cytokine-like effects through binding to various receptors including the receptor for advanced glycation end products (RAGE) [40]. Of all S100 proteins, S100A3 has the highest content of cysteine and the highest affinity for Zn^{2+}. It is expressed in human hair cuticle where it is citrullinated and tetramerized to improve its Ca^{2+}-binding ability and it contributes to the endocuticle rigidity of aged hair [41]. S100A13 on the other hand is associated with a non-classical pathway of IL-1α and FGF-1 secretion [42, 43]. It was demonstrated to form a heterotetrameric complex with IL-1α [43].

In this study, we provide evidence implicating S100A3 and S100A13 in pulmonary fibrosis. We describe an atypical form of lung fibrosis radiologically characterized by traction bronchiectasis, intralobular septal thickening and shrinkage of secondary pulmonary lobules mainly around the bronchovascular bundle with a bronchocentric interstitial fibrosis. The disease is early onset leading to type 2 respiratory failure likely attributed to the bronchocentric nature of the fibrosis around major airway causing airflow limitation. Histological analysis demonstrated advanced fibrosis and microscopic honeycombing sparing the subpleural area. The disease segregates in patients from two unrelated families in which digenic mutations in *S100A3* and *S100A13* lead to significant reduction in the expression of both proteins

The mutation in *S100A3* introduces an eleventh cysteine residue which replaces arginine residue at position 77. The p.R77C mutation is classified as a SNP having an allele frequency 0.0009635

(267/277116) and has been found in the homozygous state (4 homozygous in South Asian population) in the gnomAD database (http://gnomad-old.broadinstitute.org/variant/1-153520235-G-A). Interestingly the S100A3 arginine residue replaced in the patients is completely conserved among orthologs in five mammalian species and it is located within one of the two EF-calcium binding motifs of the protein, suggesting more severe consequences for the protein function than was predicted by the Polyphen-2 and SIFT programs. The frameshift mutation of *S100A13* is expected to result in a truncated form of the protein. Sequencing of the full intronic regions and 3' untranslated region of *S100A3* and *S100A13* excluded any other disease-associated variations of *S100A3* or *S100A13* in affected members of both families. Interestingly, sequencing of the full coding region of both genes in 28 patients with sporadic IPF did not reveal any mutations in either of these genes. However, no other cases of familial pulmonary fibrosis were genetically analyzed. The mutations in families from this study did not alter any predicted microRNA binding sites that could affect protein expression.

While *in silico* and functional evidence clearly support pathogenicity associated with co-inheritance of both variants, it remains unclear if either variant alone is sufficient to cause disease. Population data would suggest that the *S100A3* variant, while rare, may be too frequent in normal individuals to independently cause disease. Conversely, no population data exists for the *S100A13* variant and there are no instances to date of symptomatic or asymptomatic individuals with null mutations of *S100A13*. As this variant has to date only been observed in strong linkage disequilibrium with an essentially null *S100A3* variant, it is not possible to know if it independently causes disease. However, the question arises as to a possible modulating role of *S100A3* variant on the expression of S100A13. Given the *S100A13* variant is a frameshift mutation resulting in premature truncation and loss of function, any modulating role of S100A3 if present is unlikely to contribute to disease predisposition in this instance. Indeed, population frequency of the *S100A3* variant argues against it being pathogenic. Accordingly, it is unlikely that it impacts *S100A13* expression in a deleterious manner. It is possible though, that the activity of S100A3 is complemented by S100A13 and *vice vers*a, (this is particularly pertinent since both proteins are involved in the calcium signaling pathway), which might explain the digenic inheritance of the disease in the families studied.

Both mutations result in a significant reduction in the levels of corresponding proteins, which is paralleled by reduced mRNA levels. The underlying mechanism(s) behind the reduced mRNA levels is yet to be determined. Whereas the reduced *S100A13* mRNA levels due to frameshift mutation might be explained by post transcriptional nonsense mediated mRNA decay, the reduced *S100A3* mRNA levels is difficult to reconcile especially since *in silico* investigations predicted no effect of the mutation on predicted microRNA binding sites. On the other hand, a possible potential alteration of splicing mechanism(s)

which was predicted by Human Splicing Finder may account for the low mRNA levels seen in our patients' samples. This is yet to be confirmed.

Appropriate intracellular calcium homeostasis is paramount to certain stimulus response coupling, in which engagement of cognate receptors evokes the correct cellular response. Inappropriate signaling can lead to aberrant cellular behavior underlying many diseases. Here we demonstrate that the digenic mutation is associated with aberrant calcium changes in response to two independent agonists bradykinin and FGF-2, suggesting a pivotal role for S100A3/S100A13 in receptor-induced calcium transients. Direct effect of S100 proteins on intracellular calcium is not uncommon. It is noteworthy that S100A1 regulates voltage sensing and calcium release necessary for excitation contraction coupling in mouse skeletal muscle [44]. Together with calmodulin, S100A1 has been shown to bind and modulate ryanodine receptor (RyR) -dependent calcium release [45]. In addition, action potential-induced calcium transients are suppressed in skeletal muscle cells from S100A1 knockout mice [46]. Whether S100A3/S100A13 proteins behave in a similar way to S100A1 is yet to be determined.

Central to receptor mediated calcium transients is mitochondrial calcium uptake which plays a crucial role in $[Ca^{2+}]_i$ signaling by shaping and buffering calcium transients [30]. Here we demonstrate that cells isolated from patients carrying the *S100A3/S100A13* mutations display increased numbers of mitochondria and an aberrant response to oxidative challenge with signs of autophagy, which might underpin the aberrant calcium response seen in the patient fibroblasts. Moreover, direct measurement of mitochondrial Ca^{++} showed a significantly different response to bradykinin stimulation between patient and control fibroblasts.

IPF has been shown to be associated with increased levels of type I collagen and elevated levels of MMPs in blood and lung samples [31]. Since MMP expression, structure, and activity are modulated by calcium, we measured MMP levels in patient-derived cells. Our study demonstrates elevated levels of MMP1, MMP2, MMP3, MMP9 and MMP14 and reduced levels of TIMP1, which were paralleled by differential expression of collagens and other ECM-containing proteins in cultured cells isolated from our patients. Finally, in this study we demonstrate for the first time, a co-inheritance of two mutations leading to diminished expression of two calcium binding proteins in patients with an atypical form of lung fibrosis. We provide evidence that altered calcium homeostasis is associated with differential expression of ECM components in cells isolated from patients with PF. A direct link between calcium homeostasis and lung function was also reported in bleomycin models of lung fibrosis [47]. Furthermore, agonist-induced calcium oscillations have been demonstrated in pulmonary fibroblasts where they are hypothesized to control fibrotic gene expression [48]. Our data lend support to this hypothesis and suggest a "calcium-based therapy" to be considered for the management of this as yet incurable disease.

Author contributions

EAA, SPA, JH, FI, MK and FA
Conceptualization, Investigation, Resources, supervision and writing
SA, SS, ML, MA, LA, AA, AMA, MS, AM, MAC, IGL, MAO, KR, WI, AAb, FI, EAA, KSK and FA
Investigation, Data curation, formal analysis and methodology
JLJ, BFM, WC, AAl, KSC, SPA, KSK, JH, KR, FI, EAA and FA.
Writing Review and editing

Acknowledgments

We are thankful to Dr. Rachid Sougrat (KAUST) for performing TEM, Dr. Mohammed Akhtar for help with histology images, Mr. Amer Almzroua for flow cytometry experiments, Ms Turfah AlSheddi and Mr. John Scheider for technical assistance. The authors are indebted to the Saudi Genome Project (King AbdulAziz City for Science and Technology), Research Centre Administration, Training and Education Office and Office of Research Affairs, KFSHRC.

References

1. Nogee, L.M., *Genetics of pediatric interstitial lung disease.* Current opinion in pediatrics 2006. **18**: p. 287-92.
2. Raghu, G., et al., *An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management.* Am J Respir Crit Care Med, 2011. **183**(6): p. 788-824.
3. Yusen, R.D., et al., *The Registry of the International Society for Heart and Lung Transplantation: Thirtieth Adult Lung and Heart-Lung Transplant Report--2013; focus theme: age.* J Heart Lung Transplant, 2013. **32**(10): p. 965-78.
4. Marshall, R.P., R.J. McAnulty, and G.J. Laurent, *The pathogenesis of pulmonary fibrosis: is there a fibrosis gene?* Int J Biochem Cell Biol, 1997. **29**(1): p. 107-20.
5. Lawson, W.E., et al., *Genetic mutations in surfactant protein C are a rare cause of sporadic cases of IPF.* Thorax, 2004. **59**(11): p. 977-80.
6. Wang, Y., et al., *Genetic defects in surfactant protein A2 are associated with pulmonary fibrosis and lung cancer.* Am J Hum Genet, 2009. **84**(1): p. 52-9.
7. Seibold, M.A., et al., *A common MUC5B promoter polymorphism and pulmonary fibrosis.* N Engl J Med, 2011. **364**(16): p. 1503-12.
8. Kropski, J.A., et al., *Genetic studies provide clues on the pathogenesis of idiopathic pulmonary fibrosis.* Dis Model Mech, 2013. **6**(1): p. 9-17.
9. Tsakiri, K.D., et al., *Adult-onset pulmonary fibrosis caused by mutations in telomerase.* Proc Natl Acad Sci U S A, 2007. **104**(18): p. 7552-7.
10. Armanios, M.Y., et al., *Telomerase mutations in families with idiopathic pulmonary fibrosis.* N Engl J Med, 2007. **356**(13): p. 1317-26.
11. Cronkhite, J.T., et al., *Telomere shortening in familial and sporadic pulmonary fibrosis.* Am J Respir Crit Care Med, 2008. **178**(7): p. 729-37.
12. Plantier, L., et al., *Physiology of the lung in idiopathic pulmonary fibrosis.* Eur Respir Rev, 2018. **27**(147).
13. Lindner, T.H. and K. Hoffmann, *easyLINKAGE: a PERL script for easy and automated two-/multi-point linkage analyses.* Bioinformatics, 2005. **21**(3): p. 405-7.

14. Li, H. and R. Durbin, *Fast and accurate long-read alignment with Burrows-Wheeler transform.* Bioinformatics, 2010. **26**(5): p. 589-95.
15. McKenna, A., et al., *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data.* Genome Res, 2010. **20**(9): p. 1297-303.
16. DePristo, M.A., et al., *A framework for variation discovery and genotyping using next-generation DNA sequencing data.* Nat Genet, 2011. **43**(5): p. 491-8.
17. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
18. Hubbard, T., et al., *The Ensembl genome database project.* Nucleic Acids Res, 2002. **30**(1): p. 38-41.
19. Shamseldin, H.E., et al., *A null mutation in MICU2 causes abnormal mitochondrial calcium homeostasis and a severe neurodevelopmental disorder.* Brain, 2017. **140**(11): p. 2806-2813.
20. Alaiya, A.A., et al., *Protein signatures as potential surrogate biomarkers for stratification and prediction of treatment response in chronic myeloid leukemia patients.* Int J Oncol, 2016. **49**(3): p. 913-33.
21. Colak, D., et al., *Integrated Left Ventricular Global Transcriptome and Proteome Profiling in Human End-Stage Dilated Cardiomyopathy.* PLoS One, 2016. **11**(10): p. e0162669.
22. Alaiya, A., et al., *Proteomic analysis of soft tissue tumor implants treated with a novel polybisphosphonate.* Cancer Genomics Proteomics, 2014. **11**(1): p. 39-49.
23. Al-Moghrabi, N., et al., *The molecular significance of methylated BRCA1 promoter in white blood cells of cancer-free females.* BMC Cancer, 2014. **14**: p. 830.
24. Al-Mohanna, F.A., K.W. Caddy, and S.R. Bolsover, *The nucleus is insulated from large cytosolic calcium ion changes.* Nature, 1994. **367**(6465): p. 745-50.
25. O'Driscoll, B.R., et al., *BTS guideline for oxygen use in adults in healthcare and emergency settings.* Thorax, 2017. **72**(Suppl 1): p. ii1-ii90.
26. Graveley, B.R., *Sorting out the complexity of SR protein functions.* Rna, 2000. **6**(9): p. 1197-211.
27. Cartegni, L., S.L. Chew, and A.R. Krainer, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. **3**(4): p. 285-98.
28. Black, D.L., *Mechanisms of alternative pre-messenger RNA splicing.* Annu Rev Biochem, 2003. **72**: p. 291-336.
29. Bernardi, P.R., A.;, *Calcium and Cell Death: The mitochondrial connection*, in *Calcium signalling and disease: Molecula pathology of calcium.*, E.B. Carafoli, Marisa,;, Editor. 2007, Springer: New York. p. 481-506.
30. Contreras, L., et al., *Mitochondria: the calcium connection.* Biochim Biophys Acta, 2010. **1797**(6-7): p. 607-18.
31. Craig, V.J., et al., *Matrix metalloproteinases as therapeutic targets for idiopathic pulmonary fibrosis.* Am J Respir Cell Mol Biol, 2015. **53**(5): p. 585-600.
32. Hamai, K., et al., *Comparative Study of Circulating MMP-7, CCL18, KL-6, SP-A, and SP-D as Disease Markers of Idiopathic Pulmonary Fibrosis.* Dis Markers, 2016. **2016**: p. 4759040.
33. Deloulme, J.C., et al., *Nuclear expression of S100B in oligodendrocyte progenitor cells correlates with differentiation toward the oligodendroglial lineage and modulates oligodendrocytes maturation.* Mol Cell Neurosci, 2004. **27**(4): p. 453-65.
34. Halawi, A., O. Abbas, and M. Mahalingam, *S100 proteins and the skin: a review.* J Eur Acad Dermatol Venereol, 2014. **28**(4): p. 405-14.
35. Amin, A.R. and A.B. Islam, *Genomic analysis and differential expression of HMG and S100A family in human arthritis: upregulated expression of chemokines, IL-8 and nitric oxide by HMGB1.* DNA Cell Biol, 2014. **33**(8): p. 550-65.

36. Bertheloot, D. and E. Latz, *HMGB1, IL-1alpha, IL-33 and S100 proteins: dual-function alarmins.* Cell Mol Immunol, 2017. **14**(1): p. 43-64.
37. Kuberappa, P.H., et al., *Certainty of S100 from Physiology to Pathology.* J Clin Diagn Res, 2016. **10**(6): p. ZE10-5.
38. Santamaria-Kisiel, L., A.C. Rintala-Dempsey, and G.S. Shaw, *Calcium-dependent and -independent interactions of the S100 protein family.* Biochem J, 2006. **396**(2): p. 201-14.
39. Schafer, B.W. and C.W. Heizmann, *The S100 family of EF-hand calcium-binding proteins: functions and pathology.* Trends Biochem Sci, 1996. **21**(4): p. 134-40.
40. Heizmann, C.W.A., G.E.; Calichet, A., *Pathologies involving the S100 proteins and RAGE*, in *Calcium Signalling and Disease: Molecular Pathology of Calcium. Subcellular Biochemistry*, E.C.a.M. Brini, Editor. 2007, Springer: New York.
41. Takahashi, T., et al., *Age-dependent damage of hair cuticle: contribution of S100A3 protein and its citrullination.* J Cosmet Dermatol, 2016. **15**(3): p. 211-8.
42. Landriscina, M., et al., *Copper induces the assembly of a multiprotein aggregate implicated in the release of fibroblast growth factor 1 in response to stress.* J Biol Chem, 2001. **276**(27): p. 25549-57.
43. Mohan, S.K. and C. Yu, *The IL1alpha-S100A13 heterotetrameric complex structure: a component in the non-classical pathway for interleukin 1alpha secretion.* J Biol Chem, 2011. **286**(16): p. 14608-17.
44. Prosser, B.L., et al., *The Qgamma component of intra-membrane charge movement is present in mammalian muscle fibres, but suppressed in the absence of S100A1.* J Physiol, 2009. **587**(Pt 18): p. 4523-41.
45. Prosser, B.L., E.O. Hernandez-Ochoa, and M.F. Schneider, *S100A1 and calmodulin regulation of ryanodine receptor in striated muscle.* Cell Calcium, 2011. **50**(4): p. 323-31.
46. Prosser, B.L., et al., *S100A1 promotes action potential-initiated calcium release flux and force production in skeletal muscle.* Am J Physiol Cell Physiol, 2010. **299**(5): p. C891-902.
47. Mukherjee, S., et al., *Disruption of Calcium Signaling in Fibroblasts and Attenuation of Bleomycin-Induced Fibrosis by Nifedipine.* Am J Respir Cell Mol Biol, 2015. **53**(4): p. 450-8.
48. Janssen, L.J., S. Mukherjee, and K. Ask, *Calcium Homeostasis and Ionic Mechanisms in Pulmonary Fibroblasts.* Am J Respir Cell Mol Biol, 2015. **53**(2): p. 135-48.

Abbreviations

Pulmonary fibrosis (PF), Idiopathic pulmonary fibrosis (IPF), interstitial lung diseases (ILDs), runs of homozygosity (ROH), reactive oxygen species (ROS), intracellular calcium $[Ca^{++}]_i$, 4-(6Acetoxymethoxy-2,7-difluoro-3-oxo-9-xanthenyl)-4′-methyl-2,2′-(ethylenedioxy)dianiline-N,N,N,N′tetraacetic acid tetrakis (acetoxymethyl) ester (Fluo 4-AM), matrix metalloproteinase (MMP), tissue inhibitor of metalloproteinase (TIMP). Shared variant network (SVN).

Figure Legends

Figure 1: Clinical characteristics of affected patients with pulmonary fibrosis

(A) Shows the pedigrees of all families with PF with subsequent genotype analyses. Arrows indicate the proband from each family. Open symbols, members not included in the study, open symbols with genotype; unaffected; filled black symbols PF affected; (+) wild type sequence of *S100A13* /wild type "C" allele of *S100A3*; (-) 4bp deletion of *S100A13* (c.238-241delATTG)/mutant "T" allele of *S100A3* (c.299C>T). (B) CT scan at initial presentation of the three patients (F1:IV-1,2 and 3) showing central traction bronchiectasis (arrow). The distribution of fibrotic changes was peri-bronchovascular. The periphery of the lungs was spared. Global volume loss was seen with retracting sub-pleural fat in the lateral portions of the fissures (arrow heads). (C) CT scans late in the disease course of patient F1:IV-2; (i) Upper chest axial view (ii) Lower chest axial view and (iii) coronal view. There was progression of the peri-bronchovascular fibrotic changes and volume loss. Patches of ground glass densities were randomly distributed. (D&E) Two patients (F2:IV-7 & 8), CT high resolution scans showing central traction bronchiectasis (arrows). The distribution of fibrotic changes is peri-bronchovascular and central. The periphery of the lungs was spared. Global volume loss is seen evident by retracting sub-plural fat in the lateral portions of the fissures (arrow heads). (F) Pulmonary function test of 5 patients and family two members heterozygous for both the p.I80Gfs*13 and p.R77C variants in *S100A13* and *S100A3*, respectively. (G) Pathology of one affected patient (F1:IV-2): generalized interstitial inflammation with fibrosis. The inflammation mostly consists of lymphocytes in a back ground of moderate interstitial fibrosis. No advanced lung fibrosis with honeycombing is identified. No granulomas, microgranuloma or vasculitis is noted. (H) Pathology of another affected patient (F2:IV-7): interstitial inflammation with fibrosis in a diffuse pattern with no temporal heterogeneity. Advanced fibrosis seems to be sparing the subpleural space. No granulomas, microgranuloma or vasculitis is noted. Extensive sampling did not reveal a Usual Interstitial Pneumonia-like pattern.

Figure 2: Molecular analyses in pulmonary fibrosis of Families 1(A&B) & 2

(A) Shows a single ROH as a result of homozygosity mapping shared by all 7 affected patients between rs10802117 and rs11808053 confirming linkage analysis. In addition, a total of 24 unaffected family members displayed no homozygosity for this region of interest. (B) Shows linkage analysis using a total of 17 individuals (7 affected and 10 unaffected) from the two families resulting in a peak where the maximum multipoint LOD score was 5.28 corresponding to chromosome 1p12-q21.3. (C) Shows the sequence chromatogram indicating the wild-type, homozygous affected and heterozygous carrier forms of the c.238-241delATTG (p.I80Gfs*13) in *S100A13* and the C to T transition at position c.229 changing

arginine residue to cysteine at position 77 of the S100A3 protein (c.229C>T, p.R77C). Mutation name is based on the full-length *S100A3* (NM_002960) and *S100A13* (NM_001024210) transcripts.

Figure 3: Effect of *S100A3* and *S100A13* mutations on protein expression

Immunofluorescence micrographs demonstrating reduced expression of S100A3 and S100A13 proteins in lung tissue from a normal control and an affected family member (F1: IV-2) (upper images) (A). The squares are zoomed areas shown in the lower images. Relative protein expression of S100A3 and S100A13 in normal control and lung tissues from two independent patients are shown in (B) together with relative protein expression in an IPF patient. Histograms are mean intensity ± SD of multiple fields in the stained samples. (C) Confocal fluorescence laser scanning micrographs showing the reduced expression of S100A3 and S100A13 proteins in skin fibroblasts isolated from patients compared with control and the corresponding Western blots. Scale bar is 20µm. Data are representative of three independent experiments with cells isolated from two patients and two controls. (D) Relative mRNA expression of *S100A3* and *S100A13* in skin fibroblast isolated from normal control and patients. Data are representative of at least 3 independent experiments. *P-values* are depicted when appropriate.

Figure 4: Effect of *S100A3* and *S100A13* mutations on intracellular calcium changes

(A) Intracellular calcium changes following stimulation of cultured skin fibroblasts isolated from a healthy control (blue) or a patient (red). Cells were stimulated with bradykinin (50µM) (arrow). (B) The histograms show maximum response to bradykinin. Experiments were performed on live single cells using confocal laser scanning microscopy. Data are expressed as mean ± SEM (n=23 and 24 for control and patients cells). Data are expressed as normalized fluorescence intensity ratio relative to the averaged three images obtained prior to the addition of stimulus and are representative of three independent experiments. Similarly, FGF-2 (10ng/ml) stimulated cells are shown in (C) & (D). Data are expressed as mean ± SEM (n=10 and 7 cells for control (blue) and patient (red). Data are representative of 116 cells and 102 cells used in 8 and 12 independent experiments from patients and control fibroblasts respectively. The histogram depicts the maximum changes evoked by FGF-2. (E) Relative maximum calcium response to ionomycin (2µM) in skin fibroblasts from controls and patients. Data are expressed as mean ± SEM n= 28, 17, for control, patient, cells, respectively. (F) Mitochondrial calcium changes following stimulation of skin fibroblasts isolated from a healthy control (blue) or a patient (red), with bradykinin (50µM). Arrow indicates addition of bradykinin. Experiments were performed in live single cells using confocal laser scanning microscopy. All data presented above are representative of cells isolated from two patients from the two unrelated families and two controls.

Figure 5: Effect of *S100A3* and *S100A13* mutations on mitochondria

(A) Confocal fluorescence micrographs of isolated skin fibroblasts labeled with Mito Tracker® Red CMXRos (2μM) and the corresponding 3D intensity maps color coded so that warm colors indicate high intensity and cold colors indicate low intensity. (B) Flow cytometry of skin fibroblasts isolated from patients and control cells stained with Mitotracker Green. The inset shows mean ± SEM of fluorescence intensity in patients and control cells. Data were performed in triplicate and are representative of at least three independent experiments using 10^6 cells/sample. (C) TEM scans of cells isolated from healthy control (left panel) and patient's cells (right panel) depicting differences in mitochondrial size (arrows) and loss of cristae. (D) Effect of externally added oxidative insult (H_2O_2 , 0.03%, arrow) on patients and control cells labeled with Mito Tracker® Red CMXRos. Data are representative of three independent experiments.

Figure 6: Effect of *S100A3* and *S100A13* mutations on ECM components

(A) Western blots of MMP2, MMP9 and TIMP1 expression by skin fibroblasts isolated from healthy controls and patient. Relative expression is depicted in the accompanying histograms. (B) Differential expression of matrixins (MMP1,3 and 14), and (C) ECM associated proteins collagen alpha-1(VI) chain (COL6A1), collagen alpha-2(I) chain (COL1A2), collagen triple helix repeat-containing protein 1 (CTHRC1), collagen alpha-1(VIII) chain (COL8A1), collagen alpha-2(VI) chain (COL6A2), and procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 (PLOD1). In B and C normalized protein abundance of significantly differentially expressed proteins between patient and control samples are shown (a fold change >1.5, and at FDR~3%). Yeast alcohol dehydrogenase standard (P00330) at a concentration of 200 fmol per injection was used for 'Hi3' absolute quantifications of all identified proteins. The histogram bars correspond to the average protein expressions between the two sample groups using label-free LC/MS expression analysis platform on the Progenesis QIfp (Nonlinear Dynamics/Waters). Data are expressed as mean ± SEM (n=3). *P values* are indicated on top of the black horizontal bars.

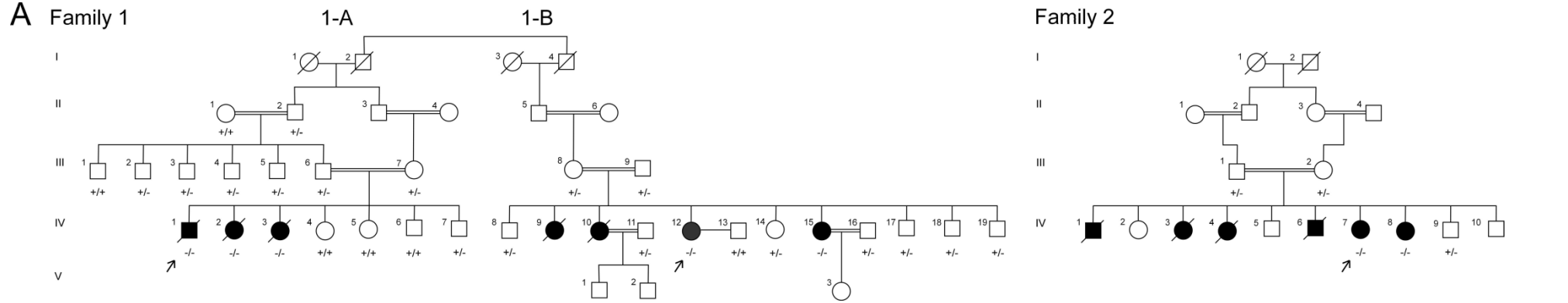
Supplementary Figure 1:

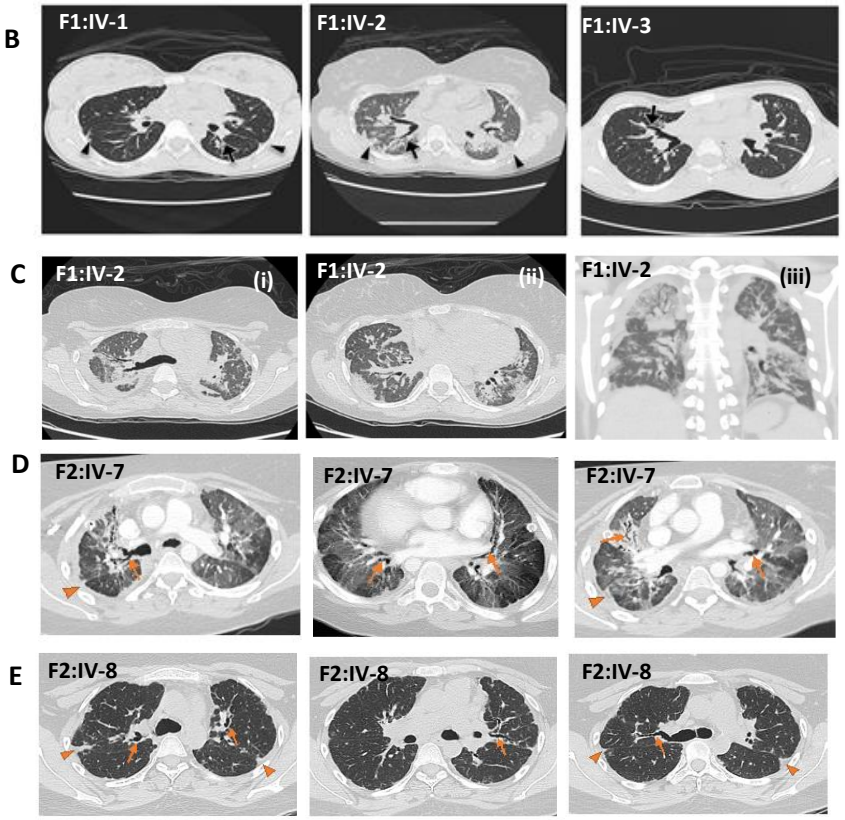
Haplotype Analysis

(A) The pedigree of families 1&2 with haplotype and genotype analyses; disease haplotype is highlighted in yellow. Genotype of variants in genes *S100A3*: c.229 C>T transition (rs138355706) causing p.R77C and *S100A13*: c.238-241delATTG causing p.I80Gfs*13 is denoted in red text. (B) Circular ideogram of the genome (Agile MultiIdeogram; http://dna.leeds.ac.uk/) showing the exclusive region of homozygosity (ROH) between the affected individuals in all families (dark blue) on Chromosome 1 (chr1:120,127,864-154,749,047 bp ; UCSC genome browser, build hg19), that is not shared with any of the unaffected

individuals (pink). Light blue and pink blocks denote ROHs present in affected and unaffected members, respectively.

Figure 1





F

Subject	FVC [L] (% Predicted)	FEV1 [L] (% Predicted)	Ratio (FEV1/FVC)	TLC [L] (% Predicted)
F1:IV-1	0.51 (15%)	0.51 (17%)	100	1.31 (28%)
F1:IV-2	0.39 (11%)	0.36 (12%)	91	1.43 (31%)
F1:IV-3	0.87 (16%)	0.84 (19%)	96	2.24 (19%)
F2:IV-7	0.54 (14.4%)	0.52 (5.9%)	97.77	2.16 (41.8%)
F2:IV-8	1.05 (30%)	0.88 (30%)	84	1.85 (36%)
F1:III-7	2.58 (86%)	2.0 (80%)	77	4.22 (90%)
F1:IV-6	3.64 (83%)	3.27 (90%)	89	4.5 (81%)

FVC= Forced Vital Capacity; FEV1= Forced Expiratory Volume in 1 second; TLC= Total Lung Capacity

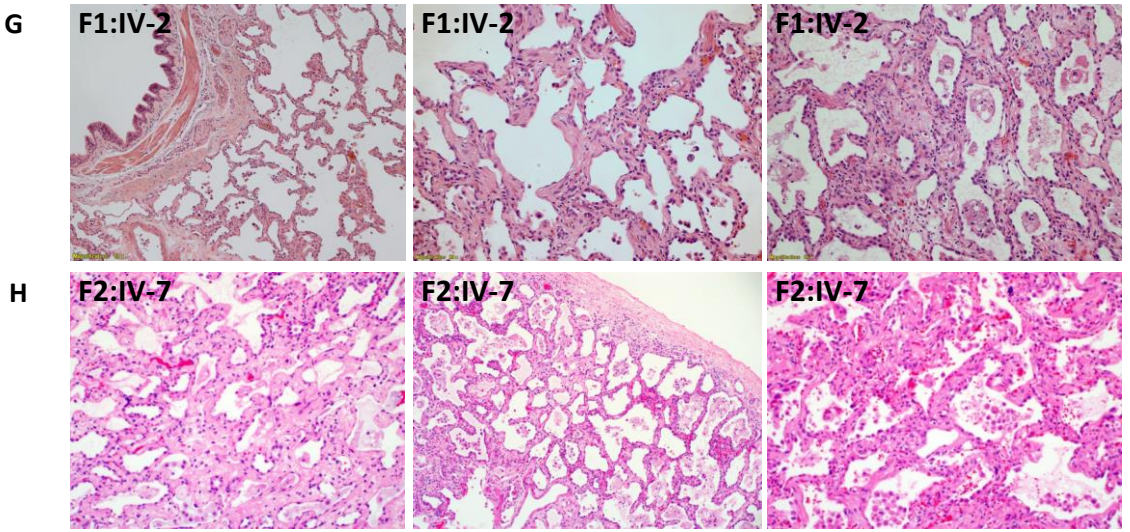
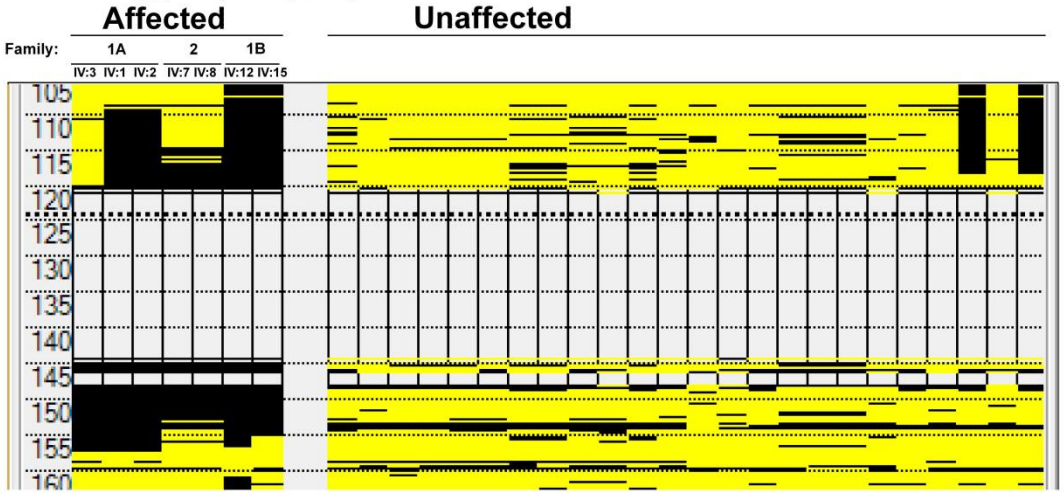
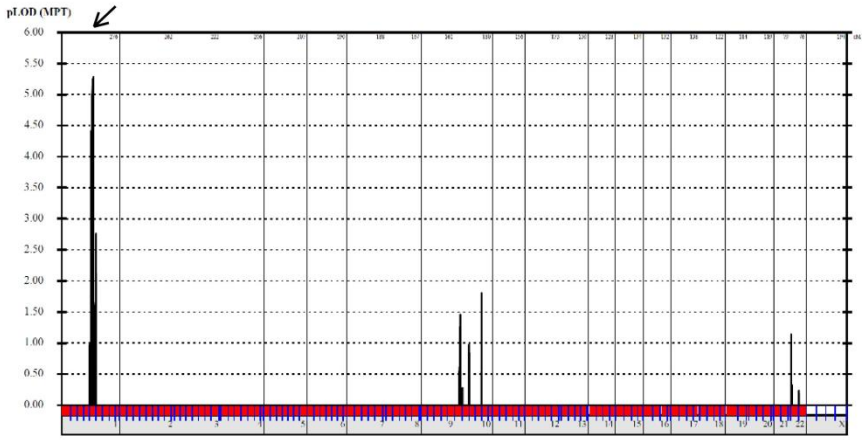


Figure 2

A Homozygosity mapping

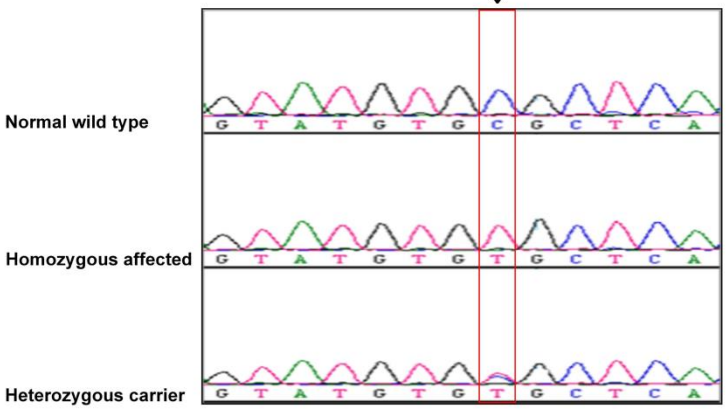


B Linkage analysis



C Sequence chromatogram

S100A3 (NM_002960)
(c.229C>T; p.R77C)



S100A13 (NM_001024210)
(c.238_241delATTG; p.I80Gfs*13)

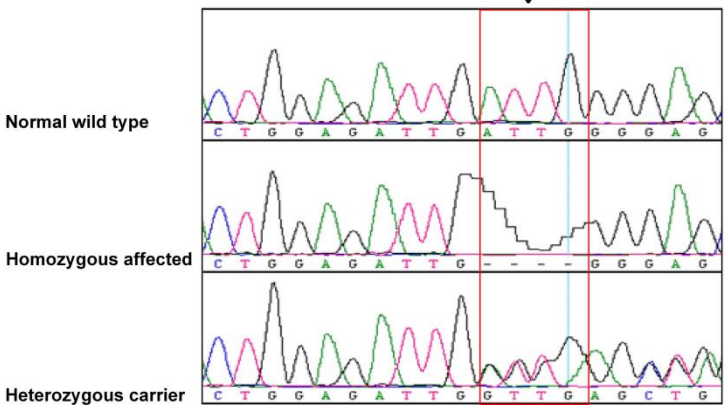
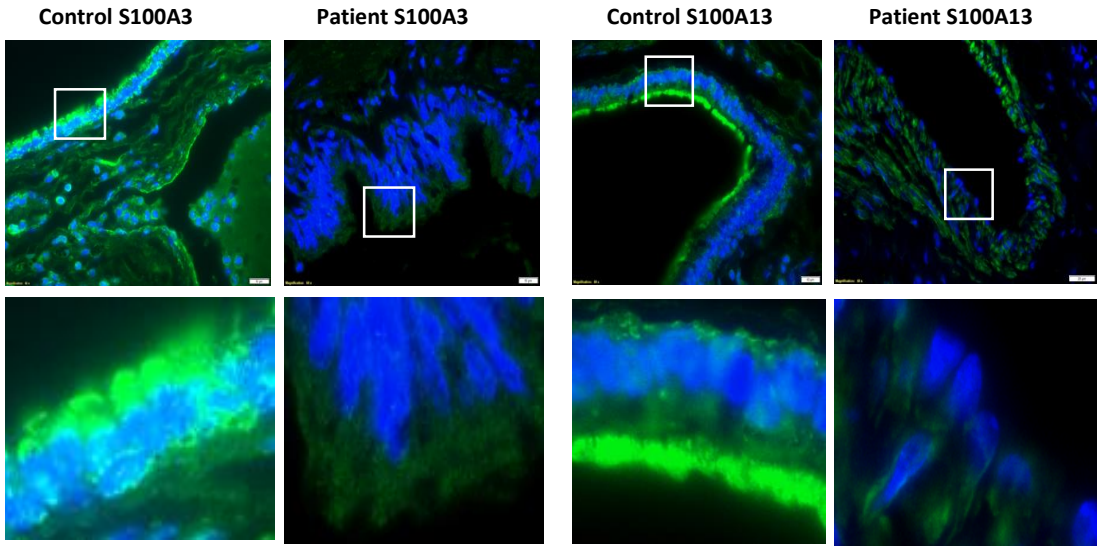
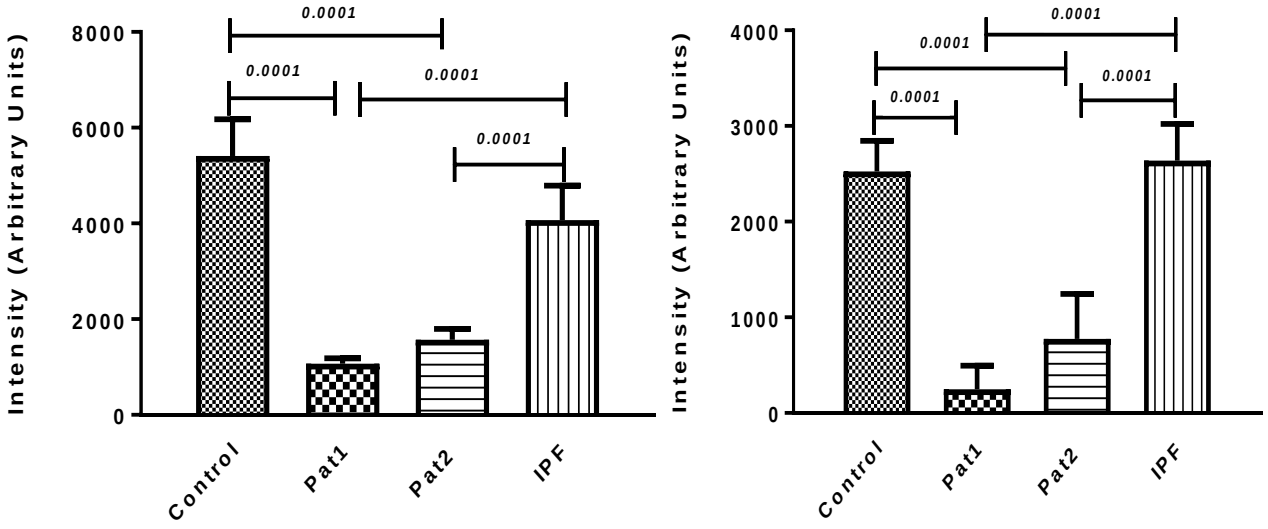


Figure 3

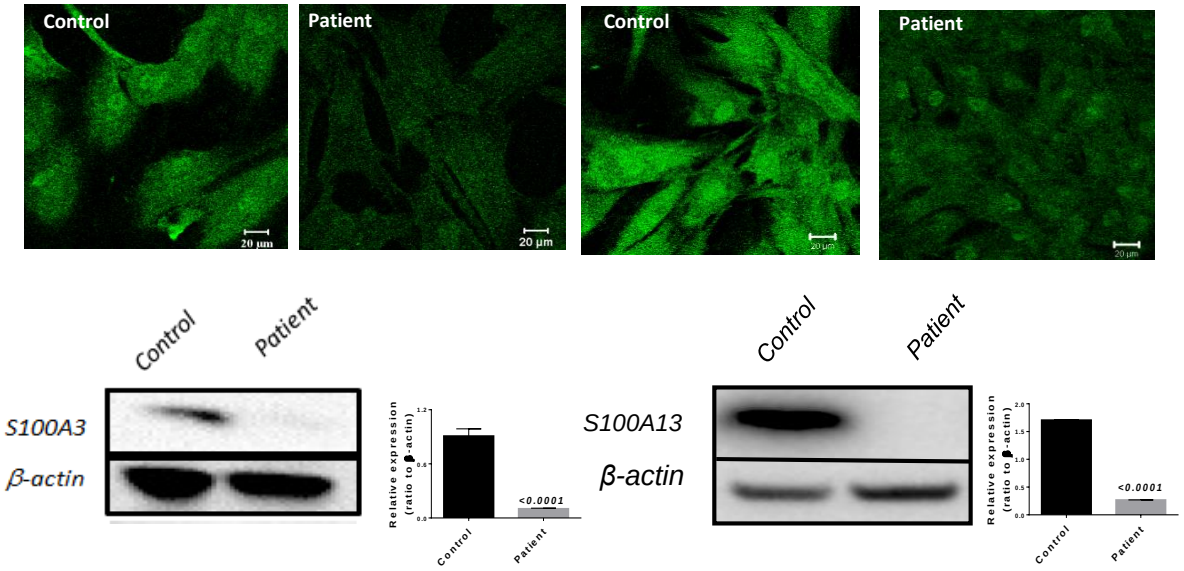
A



B



C



D

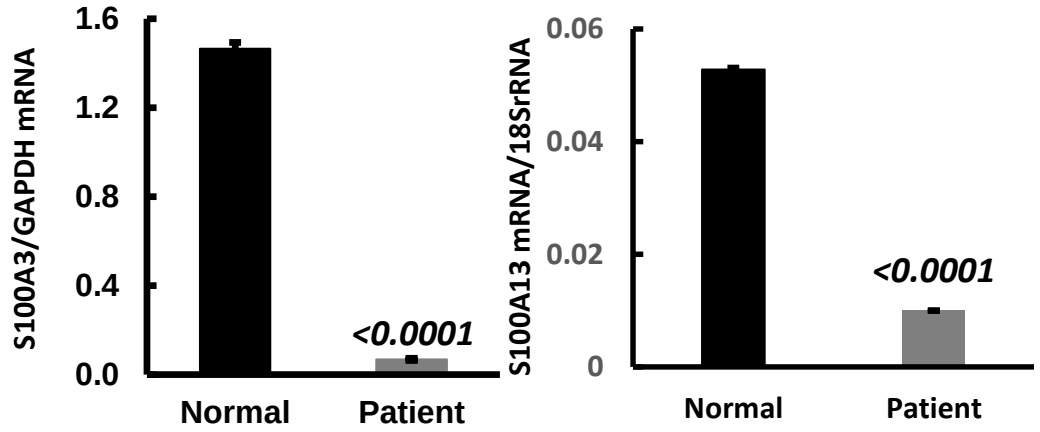


Figure 4

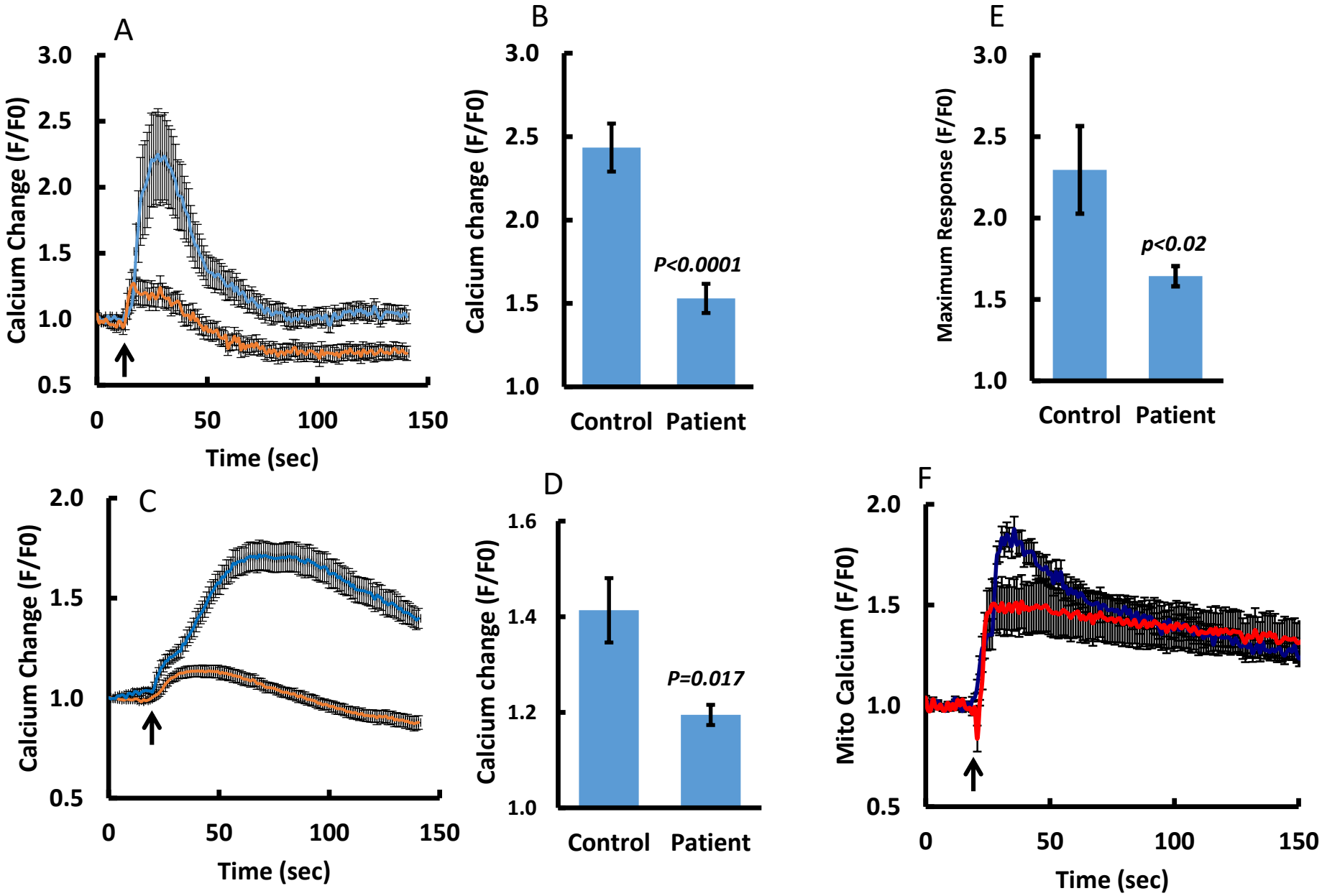

A
Calcium Change (F/F0)
3.0
2.5
2.0
1.5
1.0
0.5
0
50
100
150
Time (sec)
B
Calcium change (F/F0)
P<0.0001
Control
Patient
E
Maximum Response (F/F0)
p<0.02
Control
Patient
C
Calcium Change (F/F0)
2.0
1.5
1.0
0.5
0
50
100
150
Time (sec)
D
Calcium change (F/F0)
1.6
1.4
1.2
1.0
P=0.017
Control
Patient
F
Mito Calcium (F/F0)
2.0
1.5
1.0
0.5
0
50
100
150
Time (sec)

Figure 5

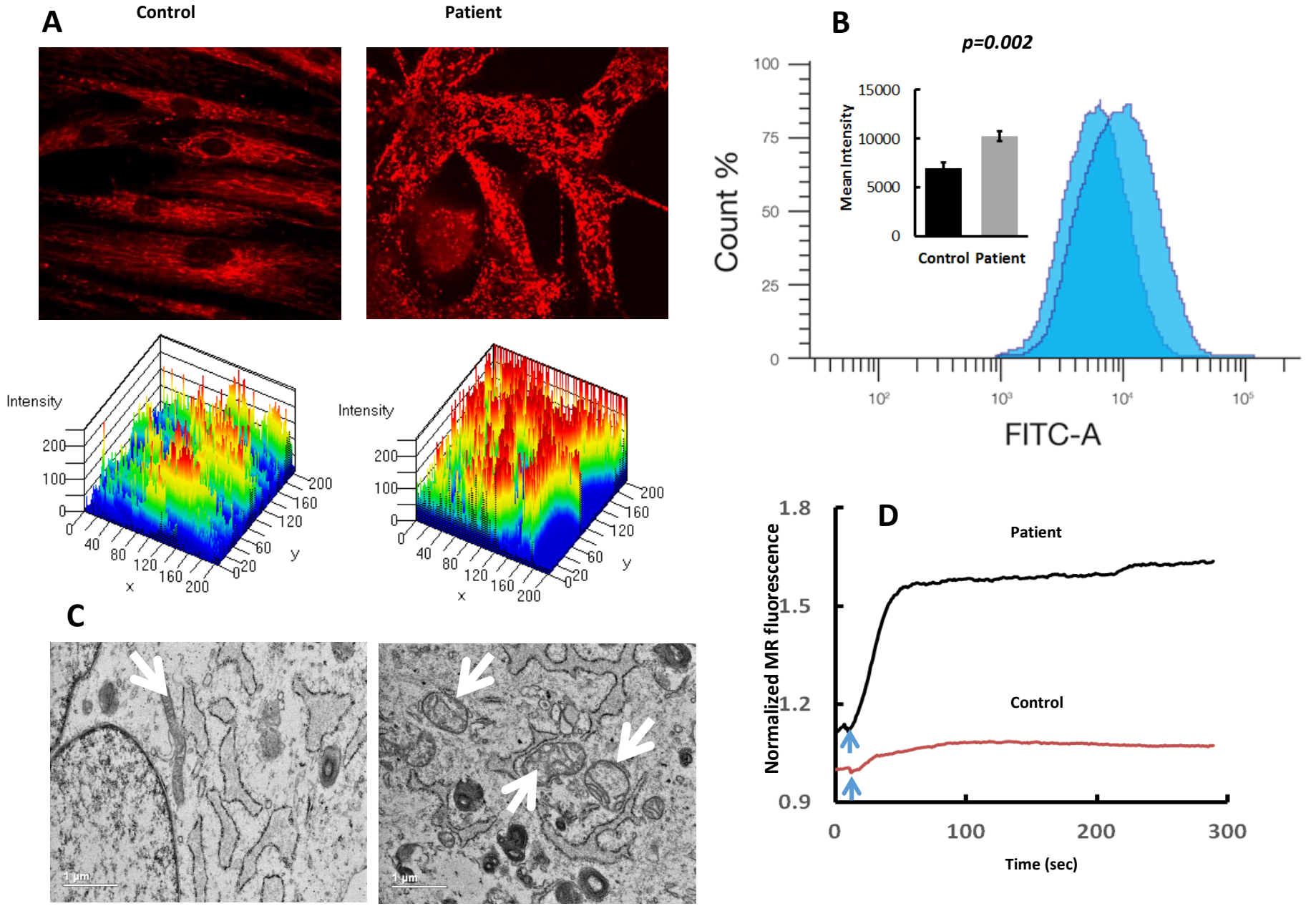
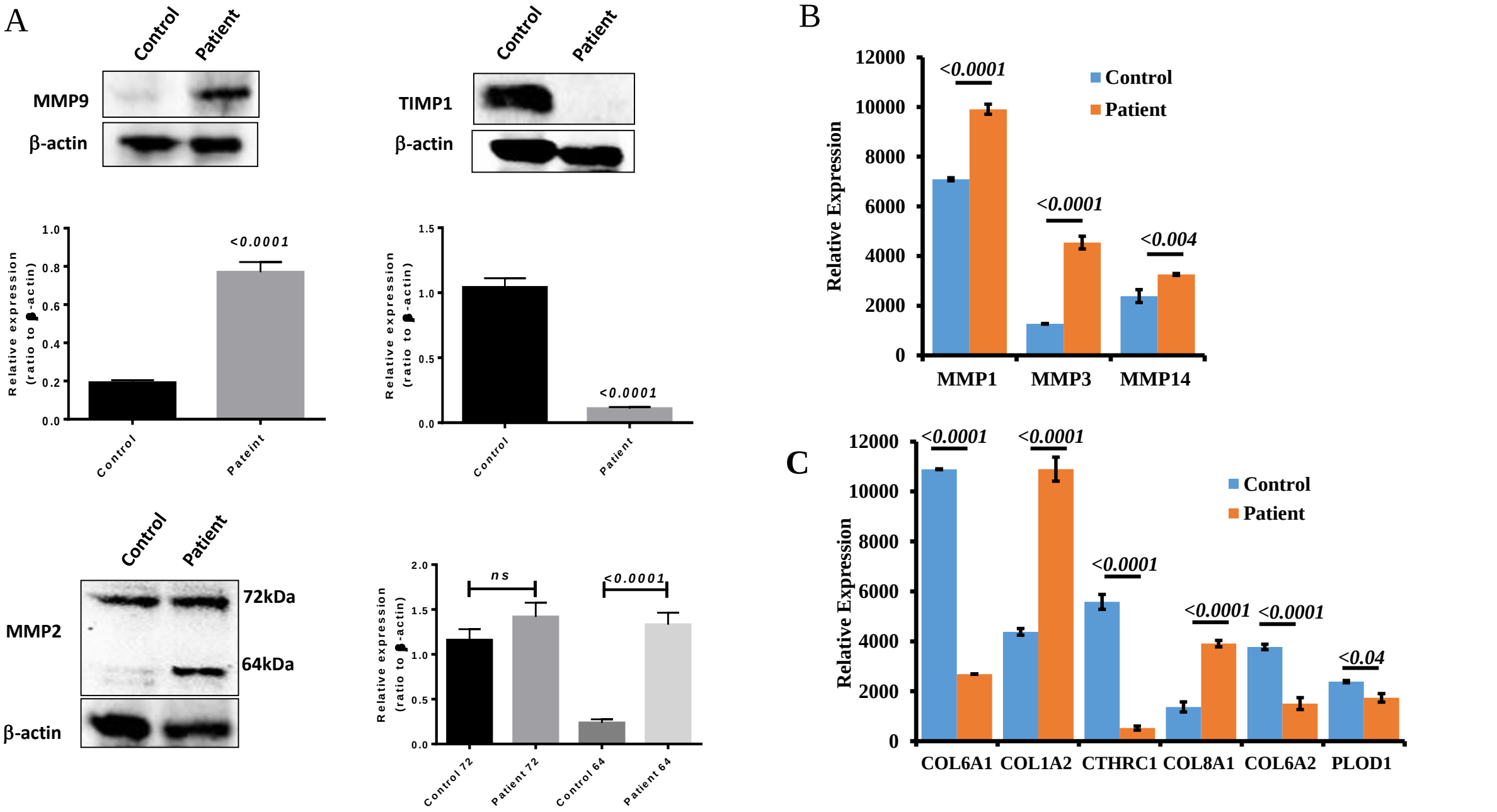
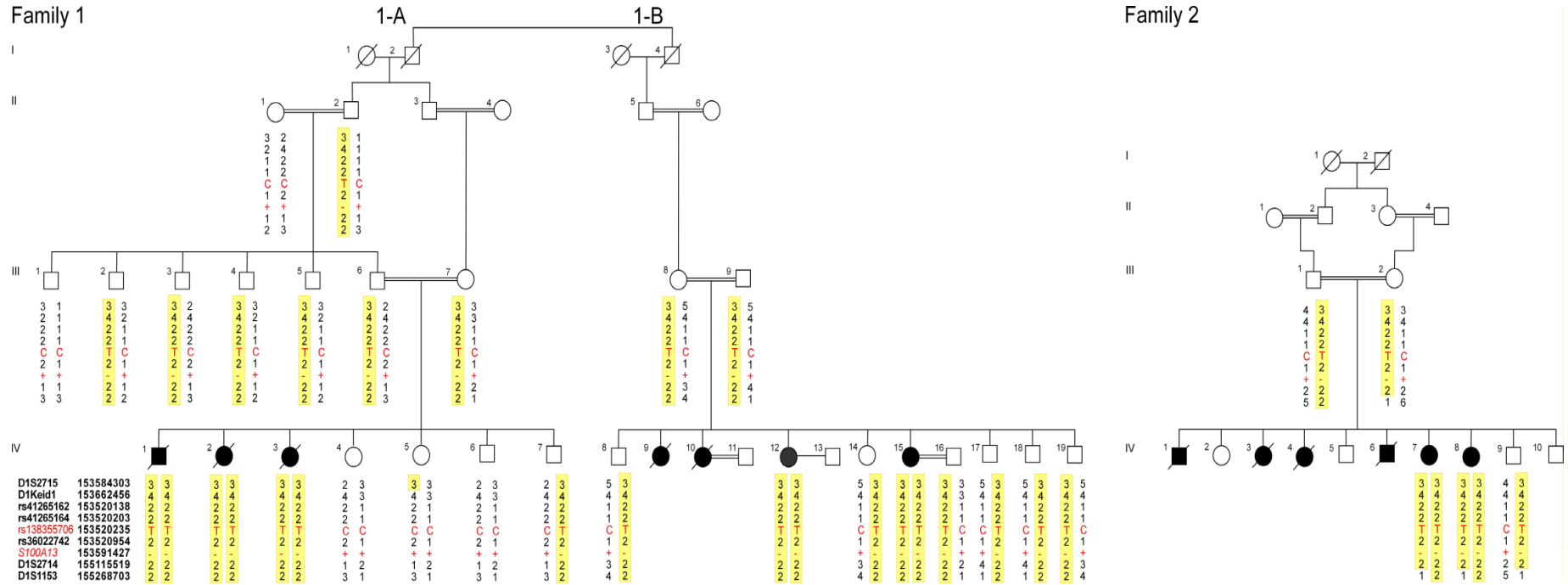

A
Control
Patient
Intensity
200
100
0
0
40
80
120
160
200
x
0
20
60
120
160
200
y
Intensity
200
100
0
0
40
80
120
160
200
x
0
20
60
120
160
200
y
B
p=0.002
100
75
50
25
0
Count %
15000
10000
5000
0
Mean Intensity
Control
Patient
10^2
10^3
10^4
10^5
FITC-A
C
1 µm
1 µm
D
Patient
Control
1.8
1.5
1.2
0.9
Normalized MR fluorescence
0
100
200
300
Time (sec)

Figure 6


A
Control
Patient
MMP9
β-actin
TIMP1
β-actin
Relative expression (ratio to β-actin)
<0.0001
Control
Pateint
Patient
MMP2
72kDa
64kDa
ns
Control 72
Patient 72
Control 64
Patient 64
B
Relative Expression
Control
Patient
MMP1
MMP3
MMP14
C
COL6A1
COL1A2
CTHRC1
COL8A1
COL6A2
PLOD1

Supplementary Figure 1A

Haplotypes of families 1&2



Supplementary Figure 1B