Πανεπιστήμιο Κρήτης
Τμήμα Ιατρικής

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Κυτταρική και Γενετική Αιτιολογία, Διαγνωστική και Θεραπευτική των Ασθενειών του Ανθρώπου»

Διδακτορική Διατριβή

ΜΕΛΕΤΗ ΤΗΣ ΠΡΩΤΕΪΝΙΚΗΣ **FLCN**

ΚΑΛΛΙΟΠΗ ΤΣΑΚΙΡΗ

Ηράκλειο, Ιούνιος 2009

Πανεπιστήμιο Κρήτης
Τμήμα Ιατρικής

Πρόγραμμα Μεταπτυχιακών Σπουδών
«Κυτταρική και Γενετική Αιτιολογία, Διαγνωστική και Θεραπευτική των Ασθενειών του Ανθρώπου»

University of Crete
Medical School

Graduate Program
Molecular Basis of Human Diseases

Διδακτορική Διατριβή

ΜΕΛΕΤΗ ΤΗΣ ΠΡΩΤΕΪΝΙΚΗΣ **FLCN**
(Γενετική της Πνευμονικής Ίνωσης)

PhD Thesis

STUDY OF PROTEIN FLCN
(Genetics of Pulmonary Fibrosis)

ΚΑΛΛΙΟΠΗ ΤΣΑΚΙΡΗ

Ηράκλειο, Ιούνιος 2009

Thesis Supervisors/ Επιβλέπων ερευνητής:
Christine Kim Garcia, MD. PHD

Supervising Commmittee/ Συμβουλευτική Επιτροπή:

D T Boumpas MD
D Georgopoulos MD
G Goulielmos PhD

Seven member committee/ Εξεταστική επιτροπή:
D T Boumpas MD
A Eliopoulos PhD
D Georgopoulos MD
G Goulielmos PhD
H Papadaki MD, PhD
K Papadakis MD, PhD
C Tsatsanis PhD

Η παρούσα εργασία πραγματοποιήθηκε / The present study was carried out at the:
McDermott Center for Human Growth and Development,
University of Texas Southwestern Medical Center,
Dallas, TX, USA

Contributions:
Christine Kim Garcia designed the research, Kalliopi Tsakiri performed research, Jennifer T. Cronkhite performed the TRFs, Melissa Nolasco provided technical support and Chao Xing assisted with the linkage analysis. Ganesh Raghu, Jonathan C. Weissler, and Randall L. Rosenblatt contributed new reagents/analytic tools. Jerry W.Shay and Wright laboratory provided assistance with telomerase assays.

OUTLINE

ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία προσδιορίσαμε μεταλλάξεις στην τελομεράση που προκαλούν ιδιοπαθή πνευμονική ίνωση των ενηλίκων, χρησιμοποιώντας συγκριτική σάρωση ολοκλήρου του γονιδιώματος με ανάλυση διασύνδεσης (whole genome linkage scan). Ελέγξαμε 46 οικογένειες με τουλάχιστο δύο μέλη προσβεβλημένα από την νόσο και 44 ασθενείς χωρίς οικογενειακό ιστορικό ιδιοπαθούς διάμεσης πνευμονίτιδας. Εντοπίσαμε επτά μεταλλάξεις στο γονίδιο *TERT* που κωδικοποιεί το πρωτεϊνικό μέρος της τελομεράσης και μια μετάλλαξη στο γονίδιο *TERC* που κωδικοποιεί το ριβονουκλεϊκό μέρος του συμπλέγματος της τελομεράσης. Πέντε από τις μεταλλάξεις στο *TERT* αφορούν σε αλλαγή βάσης και εμπλέκουν αυστηρά συντηρημένα αμινοξέα ανάμεσα στα είδη. Οι υπόλοιπες δύο μεταλλάξεις στο *TERT* αφορούν σε έλλειψη βάσεων και οδηγούν στο σχηματισμό πρόωρων κωδικονίων τερματισμού. Η μετάλλαξη που εντοπίσαμε στο *TERC* αφορά σε αλλαγή βάσης και έχει ήδη περιγραφεί σε έναν ασθενή με συγγενή δυσκεράτωση και απλαστική αναιμία. Μετρήσαμε την ενεργότητα αυτών των μεταλλαγμένων συμπλόκων αφού πρώτα εκφράσαμε τις μεταλλαγμένες πρωτεΐνες, το μεταλλαγμένο μόριο του RNA και τα αντίστοιχα αγρίου τύπου σε in vitro συνθήκες. Η ενεργότητα των μεταλλαγμένων συμπλόκων βρέθηκε να κυμαίνεται από 30 εώς 100% για τις πέντε μεταλλάξεις αλλαγής βάσεως στο *TERT* και τη μετάλλαξη στο *TERC*, ενώ για τις δύο *TERT* μεταλλάξεις έλλειψης βάσεων η ενζυμική ενεργότητα ήταν σχεδόν μη ανιχνεύσιμη. Επιπλέον, σχεδιάσαμε ένα πείραμα για να διαχωρίσουμε τον τρόπο δράσης μεταξύ επικρατούς-αρνητικής (dominant –negative) και απλοανεπάρκειας για μια από τις πιο δραματικές μας μεταλλάξεις V747fsX. Η V747fsX μετάλλαξη οδηγεί στο σχηματισμό μιας πρωτεΐνης που υπολείπεται το καρβοξυτελικό άκρο της συντηρημένης περιοχής η οποία είναι απαραίτητη για την λειτουργία της ανάστροφης μεταγραφής. Στο πείραμα αναμείξαμε σε διαφορετικές αναλογίες μεταλλαγμένη τελομεράση και τελομεράση αγρίου τύπου και μετρήσαμε την ενεργότητα των συμπλόκων που προέκυπταν. Παρατηρήσαμε ότι η αύξηση της ενεργότητας ήταν ευθέως ανάλογη με την προσθήκη του αγρίου τύπου ενζύμου υποδεικνύοντας την απλοανεπάρκεια, δηλαδή την απουσία επαρκούς ποσότητας ενζύμου, ως τον μηχανισμό δράσης της συγκεκριμένης μετάλλαξης.
Παράλληλα μετρήσαμε το μήκος των τελομεριδίων των ασθενών με ιδιοπαθή πνευμονική νόσο και των λοιπών μελών των οικογενειών τους και διαπιστώσαμε ότι οι φορείς μεταλλάξεων της τελομεράσης έχουν βραχύτερα τελομερίδια από τους μη φορείς. Αναλύοντας τα αποτελέσματα ανά οικογένεια προκύπτει ότι οι φορείς ασθενείς έχουν βραχύτερα τελομερίδια από τους φορείς μη ασθενείς και οι τελευταίοι βραχύτερα από τους μη φορείς μη ασθενείς συγγενείς τους συγκρίσιμης ηλικίας. Τέλος, υπολογίζοντας το ποσοστό των βραχύτερων τελομεριδίων ανά άτομο–καθώς τα βραχύτερα τελομερίδια καθορίζουν την μοίρα του κυττάρου και όχι το μέσο μήκος των τελομεριδίων- διαπιστώσαμε ότι οι ασθενείς φορείς έχουν το μεγαλύτερο ποσοστό βραχέων τελομεριδίων.
Από την παρούσα εργασία προκύπτουν εργαστηριακά δεδομένα βάσει των οποίων η ιδιοπαθής πνευμονική ίνωση των ενηλίκων εντάσσεται στις νόσους δυσλειτουργίας των τελομεριδίων τουλάχιστο σε ένα ποσοστό 15% των οικογενών και σποραδικών μορφών της.

IDIOPATHIC PULMONARY FIBROSIS

Idiopathic pulmonary fibrosis (IPF) is a progressive interstitial lung disease of the family of idiopathic interstitial pneumonias (IIP) (Figure 1). IIPs are a group of non neoplastic scarring lung disorders sharing clinical features and varying patterns of inflammation and fibrosis (Gross et al. 2001). The IIPs can be classified in pathologically distinct categories and Liebow 1969 was the first to divide them into five groups based on the pathologic changes observed: usual interstitial pneumonia (UIP), bronchiolitis obliterans interstitial pneumonia and diffuse alveolar damage, desquamative interstitial pneumonia, lymphocytic interstitial pneumonia, and giant cell interstitial pneumonia. Since then a number of changes have taken place by investigators in the classification of IIPs (Katzenstein et al. 1998). Although correct histological classification is very significant for the prognosis and the survival (Bjoraker et al. 1998) a minority of IIP diagnosis are made based on surgical lung biopsy. A new classification was published by the American Thoracic Society (ATS) and the European Respiratory Society (ERS) integrating clinical, radiologic and pathologic features of IIPs and the clinicopathologic criteria are shown in Table1 (ATS/ERS, 2002).

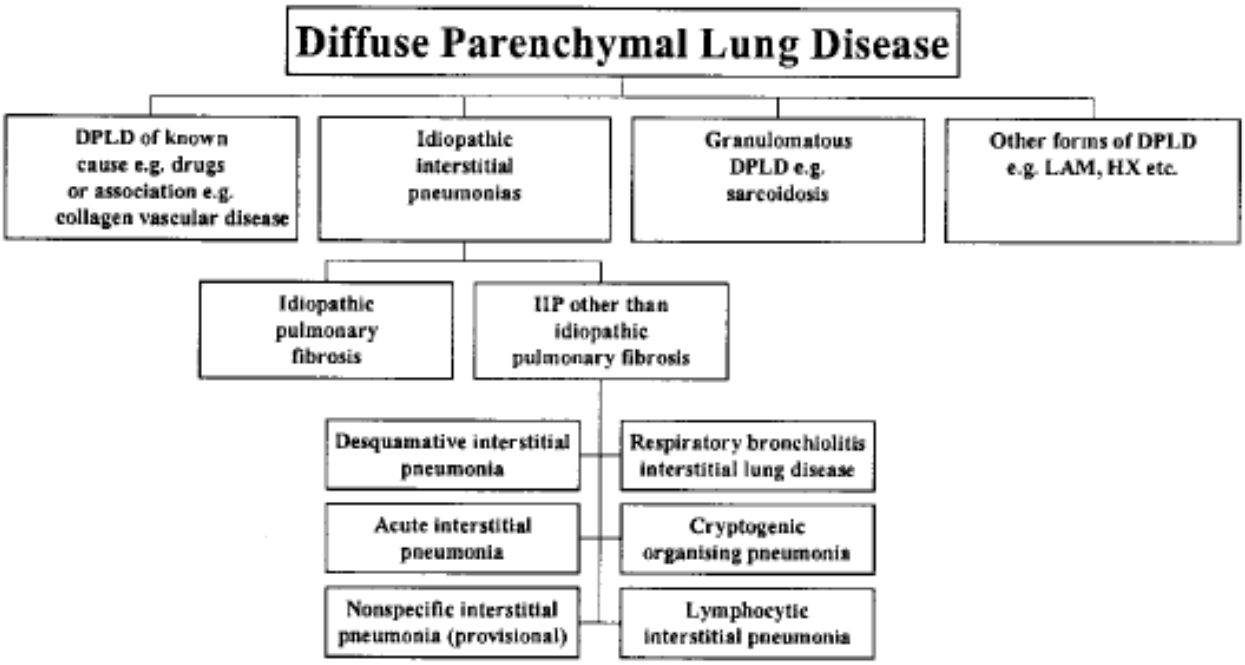

Figure 1
Diffuse parenchymal lung diseases (DPLDs) are a group of disorders of known etiology (e.g. drug toxicity –amiodarone, environmental and occupational exposures-asbestosis) or association (e.g. systemic connective tissue diseases –rheumatoid arthritis) as well as of unknown cause. Idiopathic interstitial pneumonias (IIPs), granulomatous lung disorders, and other forms of interstitial lung disease (DPLD) are included in the latter group. IIPs are divided in two groups: the idiopathic pulmonary fibrosis (IPF) and the non IPF group which include pathologic distinctive entities as shown in the diagram. LAM= lymphangioleiomyomatosis, HX= histiocytosis X. (Figure adapted from ATS/ERS, 2002).

Table 1 (adapted from ATS/ERS, 2002)

HISTOLOGIC AND CLINICAL CLASSIFICATION OF IDIOPATHIC INTERSTITIAL PNEUMONIAS*

Histologic Patterns	Clinical–Radiologic–Pathologic Diagnosis
Usual interstitial pneumonia	Idiopathic pulmonary fibrosis/cryptogenic fibrosing alveolitis
Nonspecific interstitial pneumonia	Nonspecific interstitial pneumonia (provisional)†
Organizing pneumonia	Cryptogenic organizing pneumonia
Diffuse alveolar damage	Acute interstitial pneumonia
Respiratory bronchiolitis	Respiratory bronchiolitis interstitial lung disease
Desquamative interstitial pneumonia	Desquamative interstitial pneumonia
Lymphoid interstitial pneumonia	Lymphoid interstitial pneumonia

IPF is defined as a distinctive form of chronic fibrosing interstitial pneumonia limited to the lung and associated with the histologic appearance of usual interstitial pneumonia (UIP) on surgical (thoracoscopic or open) lung biopsy. In the presence of a surgical biopsy showing a UIP pattern the definite diagnosis of IPF includes the following: 1. Exclusion of other known causes of interstitial lung disease such as drug toxicities, environmental exposures, and collagen vascular diseases 2. Characteristic abnormalities on conventional chest radiographs or high-resolution computed tomography (HRCT) scans 3. Abnormal pulmonary function studies that include evidence of restriction (reduced VC often with an increased FEV1/FVC ratio) and/or impaired gas exchange [increased AaPO2 (alveolar–arterial pressure difference for O2) with rest or exercise or decreased DLCO (diffusing capacity of the lung for CO) (ATS/ERS, 2000).

The clinical presentation of the disease is insidious with the gradual onset of a nonproductive cough, often paroxysmal and dyspnea. Dyspnea is the most prominent symptom being usually progressive and present six months before presentation. Usually the patient at the time of presentation is a middle age male adult. Constitutional symptoms including malaise, joint symptoms and weight loss are present in almost half the patients although fever is generally a rare symptom and its presence may be suggesting an alternative diagnosis. Digital clubbing develops in 25 to 50% of patients. Features of right heart failure and peripheral oedema are present only in late stages of the disease. Chest auscultation reveals fine end- inspiratory crackles to the basal areas and gradually as the disease progress involve the upper lung zones as well. Pulmonary function tests might be normal at the beginning or exhibit a restrictive pattern. The clinical course of the disease is characterised by gradual deterioration. Once the diagnosis is made median length of survivor is 2.5 to 3.5 years. Rapid decline may be due to accelerated disease, viral infection and the development of pneumonia or diffuse alveolar damage (ATS/ERS, 2002, ATS/ERS, 2000). Factors that may precipitate the death include older age (>50yr), poor pulmonary function at presentation or recent deterioration on the pulmonary tests three to six months after initial corticosteroid therapy and advanced fibrosis. Predictors of longer survival are the female sex, presence of ground glass and reticular opacities on HRCT, increased proportion of lymphocytes (20 to 25%) in BAL and being current smoker at the time of presentation (ATS/ERS, 2000). The most common cause of death is respiratory failure. Other common causes include heart failure, bronchogenic carcinoma, ischemic heart disease, infection and pulmonary embolism. The prevalence of the disease increases with the age and is more common in men compare to women. The prevalence and incidence of IPF in the United States ranked from 0.4 (age, 18-34yr) to 27.1 (age>75yr) per 100.000 (Raghu et al. 2006). 0.5 to 2.2% of all cases have a positive family history. The familial type of the disease seems to be indistinguishable from the sporadic form but for the age of onset as it

appears to be earlier in the familial form (55yr versus 68yr for the sporadic) (Marshall et al. 2000).
Non specific laboratory test abnormalities including increased erythrocyte sedimentation rate, antinuclear antibodies, cryoglobulins and rheumatoid factor may be present (Katzenstein et al. 1998). Polycythemia is rare despite the frequent presence of chronic hypoxemia (ATS/ERS, 2000).
Bronchoalveolar lavage (BAL) fluid contains an excess of neutrophils proportional to the extend of reticular changes seen on CT of the lung. A moderate increase of eosinohils might be present but when they account more than 20% eosinohilic lung disease should be considered. BAL cell counts correlate with the severity of the disease but cannot predict prognosis.
Chest radiographs show peripheral reticular opacity mostly in the bases often associated with honeycombing and lower lobe volume loss. Although occasionally chest X ray may be normal in IPF patients the CT scan in these cases will show evidence of the disease. HRCT examinations present patchy, predominantly peripheral, subpleural, bibasal reticular opacities. Honeycombing and ground glass attenuation are also common. Honeycomb cysts usually enlarge over time while architectural distortion, lobar volume loss, traction bronchiectasis and bronchiolectasis are seen in more advanced disease. HRCT allows earlier diagnosis of IPF, helps to narrow the differential diagnosis based on the CT pattern, and allows the identification of associated emphysema. The presence of typical clinical and HRCT features of UIP is sufficient in more than 50% of suspected cases to set the diagnosis. HRCT evaluation confers diagnosis of IPF in about two-thirds of patients that have histologic proven UIP. Thus about one-third of cases of UIP will be missed by relying on CT diagnosis alone (ATS/ERS, 2000).
A surgical lung biopsy is necessary for a confident clinicopathologic diagnosis of UIP/IPF when the clinico-radiologic picture is not typical. Histologic variation with alternating areas of normal lung, interstitial inflammation, fibrosis, and honeycomb change are important diagnostic clue of UIP/IPF. The peripheral subpleural parenchyma is most severely affected by these changes. Interstitial inflammation is usually mild to moderate and composed mainly of alveolar septal infiltrate of small lymphocytes, plasma cells, histiocytes and hyperplasia of type II pneumonocytes. In fact the presence of severe inflammation should suggest a diagnosis other than UIP/IPF. The fibrotic zones show temporal heterogeneity with dense acellular collagen bundles and scattered fibroblastic foci. The latter ones are small aggregates of actively proliferating myofibroblasts and fibroblasts, characterized by spindle – shaped cells and arranged parallel to the long axis of the alveolar septa. Fibroblastic foci are not pathognomonic for UIP but are necessary for the diagnosis (Figure 2). Honeycomb change are characterized by enlarged air spaces lined by bronchiolar epithelium or hyperplastic alveolar pneumocytes, and separated by thick walls containing collagen and varying amounts of chronic inflammation cells. The cystic fibrotic airspaces may be empty, or frequently contain mucin mixed with inflammatory cells. Honeycomb change is not specific to UIP and its presence implies of scarring and architectural restructuring (ATS/ERS, 2002). In the absence of lung biopsy the diagnosis is made with the clinical radiological criteria as shown in the Table 2.

Table 2 (adapted from ATS/ERS, 2002)

Criteria for the diagnosis of IPF in the absence of lung biopsy

Major Criteria
Exclusion of other known causes of interstitial lung disease
Abnormal pulmonary function studies that include evidence of restriction and impaired gas exchange
Bibasilar reticular abnormalities with minimal ground glass opacities on HRCT scans
Transbrochial lung biopsy or BAL showing no features to support an alternative diagnosis
Minor criteria
Age>50 yr
Insidious onset o otherwise unexplained dyspnea on exertion
Duration of illness> 3months
Bibasilar inspiratory crackles

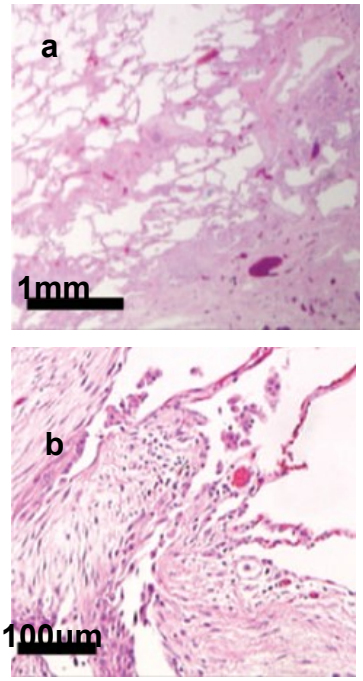

Figure 2
UIP/IPF histology pattern in lung biopsy. The thoracoscopic lung biopsy shows normal lung and densely fibrotic lung tissue with architectural distortion (a). Prominent fibroblast proliferation is more visible in a higher magnification (x10) with the presence of fibroblastic foci and loose connective tissue (b).
Figure adapted by Tsakiri et al.2007

The pathogenesis of IPF is not clear. Various hypotheses have been made about the mechanism that leads to the development of the disease. The original one suggests that pulmonary fibrosis is the response to injury from chronic inflammation. However antiflammatory therapy has offered little benefit in limiting or modifying the natural course of the disease. An alternative hypothesis implies that repeated episodes of acute lung injury result in PF. These sequential lung injury episodes in combination with aberrant wound healing ultimately lead to fibrosis and loss of lung function. Factors that are considered to modify the wound healing and the response to stimulus include environmental factors and the genetic background of the patient (Gross et al. 2001).

Smoking is considered to be an independent risk factor for the development of IPF (Steele et al. 2005) with 75% of the affected cases recorded in various studies being current or former smokers (Gross et al. 2001). Viral infections and respirable

environmental agents have been proposed to trigger the lung injury process without any convincing association.
When we started working on this study there was a lot of evidence to suggest that the development of the disease may be influenced by genetic factors. Pulmonary fibrosis is observed in genetic syndromes with pleiotropic presentation as neurofibromatosis, familial surfactant protein C mutation (SFTPC) and dyskeratosis congenita. In the case of the SFTPC a mutation was identified in kindred with many pediatric cases and incomplete penetrance. Electron Microscopy results of the affected lung tissue suggested that there is damage to the alveolar epithelial cells, especially the type II alveolar cells (Thomas et al. 2002). SFTPC mutations were detected in an infant with nonspecific interstitial pneumonitis and her mother who was diagnosed with desquamative interstitial pneumonitis at one year of age (Nogee et al. 2001). Dyskeratosis congenita (DKC) has a pleiotropic phenotype and pulmonary fibrosis is one manifestation of the syndrome. In a three generation pedigree idiopathic pulmonary fibrosis was present in four out of seven affected individuals with DKC while pulmonary fibrosis was the only manifestation of the disease for one individual (Armanios et al. 2005, 2007). Familial pulmonary fibrosis accounts for approximately 2% of the total cases of IPF (Marshall et al. 2000) and aside from these SFTPC and other genes involved in primarily pediatric disease the genetic basis of adult onset familial pulmonary fibrosis was not understood. With our study on genetics of IPF presented in the following pages we identified that mutations in telomerase and its RNA component cause adult onset pulmonary fibrosis in about 15% of familial cases and as well as in sporadic cases through telomere shortening. In this case pulmonary fibrosis may be the result of senescence of a cell population able to respond to repetitive injuries over time. Another group screened 73 probands from the Vanderbilt Familial Pulmonary Fibrosis Regisry for mutations in *TERT* and *TERC*. They found five mutations in TERT and one mutation in TERC in six probands. In addition using flow fluorescence in situ hybridization (FISH) they measured the average telomere length for the family members and found that the probands as well as the unaffected carriers had shorter telomeres compare to the non carriers of telomerase mutations (Armanios et al. 2007). This independent study confirmed the role of telomerase mutations in IPF. Recently using whole genome linkage scan it was reported that mutations in surfactant protein A2 (SFTPA2) cause pulmonary fibrosis and lung cancer. The mutations result in misfolded proteins that are retained in the endoplasmic reticulum (ER) and are not secreted. That may lead to ER stress and alveolar type II cells dysfunction (Wang, Y. et al. 2009). A plausible mechanism through which these mutations cause pulmonary fibrosis involves alveolar type II cells dysfunction and aberrant response to repetitive stimulus.
The treatment of IPF has been based on the theory that chronic inflammation is the cause of the disease. Immunosuppressive therapies have not been proven to be effective in the course of the disease. Large trials were never conducted to assess corticosteroid efficacy in pulmonary fibrosis and the evidence to support their use are based on small trials with contradictory outcomes (Meltzer and Noble. 2008). The ATS/ERS consensus statement recommends the use of prednisone in combination with azathioprine or cyclophosphamide. The duration of therapy is at least six months with clinical and physiological monitoring to asses the response to treatment (ATS/ERS, 2000). The standard therapy proposed by the ATS/ERS consensus may be complimented with N-acetylcysteine (NAC). In a double blinded study IFIGENIA (Idiopathic Pulmonary Fibrosis International Group Exploring N-Acetylcysteine I Annual) trial it was shown that prednisone/azathioprine/NAC treatment was superior

to standard therapy as the lung function deterioration was less over a 12month period without though improving or stabilizing patients with IPF (Demedts et al. 2005). N-acetylcysteine, a molecular precursor to the naturally occurring antioxidant glutathione, theoretically is targeting the oxidative injury that precedes fibroproliferation.

Another promising drug is pirfenidone an antifibrotic agent currently being tested in phase III studies (CAPACITY: Clinical Studies Assessing Perfenidone in IPF) in the United States and already approved for IPF in Japan. A study that enrolled 107 IPF patients tested pirfenidone versus placebo. The primary end point of the trial was negative but improved vital capacity and survival was to favour pirfenidone (Azuma et al. 2005).

Lung Transplantation is to be considered for patients diagnosed with IPF since it has been shown that five year survival after transplantation approaches 50-60%. Patients should be carefully selected for co morbidities since following transplantation they require life long immunosuppression. Graft rejection, infection, and heart failure are the most common causes of early mortality. Patients should be carefully listed foe transplantation taking into considerations the predictors for the patients' prognosis as well as the benefits of such an operation (Meltzer and Noble. 2008).

Identifying the pathogenesis and the mechanism that lead to pulmonary fibrosis may help develop new targeted therapies that will improve survival as well as the course of the disease.

TELOMERASE

Telomerase is a specialized reverse transcripase responsible for extending the G-rich strand of telomeres by adding tandem repeats of six base pairs TTAGGG. Telomerase is a ribonucleoprotein widely conserved among eukaryotes. In human the core enzyme consists of two main components: the telomerase reverse transcriptase (tert) and the telomerase RNA. Tert is a 1132aa protein with a predicted molecular mass of more than 100kD encoded by the *TERT* gene. The *TERT* gene is located on the short arm of chromosome 5p15.33 and consists of 16 exons and 15 introns spanning approximately 35kbs. Tert is not expressed in the majority of normal adult tissues, including cardiac and skeletal muscle, adipose tissue, lung, liver and kidney. It is expressed in cells that are considered to have long-term proliferative capacity like bone marrow cells, lymphocytes in germinal centres and spermatogonia cells in the testis as well as in immortal cancer cells (Kolquist et al. 1998). Telomerase RNA gene (*TERC*) is mapped on the long arm of chromosome 3q21-q28 and encodes a 451nts molecule that is ubiquitously expressed. The RNA molecule contains the template CAAUCCCAAUC used by telomerase to specify the repeat sequences added to telomeres. The terminal nucleotides at the telomeres are used as primers for the reverse transcription (Figure 3) (Aubert and Lansdorp. 2008). Maintenance of telomere ends by telomerase compensates for the loss of telomeric DNA during the replication process also known as "the end replication problem". The leading strand can always extend to the end because it is primed from behind. At the chromosome end though the lagging strand reaches a point where the priming system cannot work resulting in an unpolymerized section (Griffiths, 7th Edition).
Telomeres are special nucleoprotein complexes located at the end of chromosomes. Their size varies in species with human telomeres ranging from 5 -15 kbs and murine 20-150kbs. These structures prevent the end of chromosomes from being recognized as DNA break and protect the termini from fusion, degradation or recombination. Telomeres are composed by six base pairs repeats bound to a protein complex called shelterin or telosome. A number of proteins have been found to directly or indirectly associate with telosome as trf1, trf2, tin2, tpp1, rap1 and pot1. Studies have shown that these proteins have a role in telomere length regulation as well as in protection of the telomeric end from end to end fusions. Telomeric DNA typically ends in a single strand G-rich overhang of between 50 to 300nts to the 3' end of the chromosome. This overhang can fold back and form a 'T-loop' structure as a primitive way to protect the end of the chromosome. Adjacent to telomeres are the subtelomeric regions that are also rich in repetitive DNA. Telomere shortening takes place after each cell division as most human somatic cells do not have active telomerase. Human somatic cells in culture lose 50-200bps after each round of replication. Telomere length can be affected by the cellular activity of telomerase and the cell's number of cell divisions while telomere shortening can be accelerated by DNA damaging agents and nucleases. Telomerase is the main mechanism for maintaining telomere integrity. Immortal cell lines as well as cells that lack telomerase can maintain their telomere length activating the alternative lengthening of telomeres (ALT) mechanism which is based on homologous recombination (Blasco, M. A. 2007).

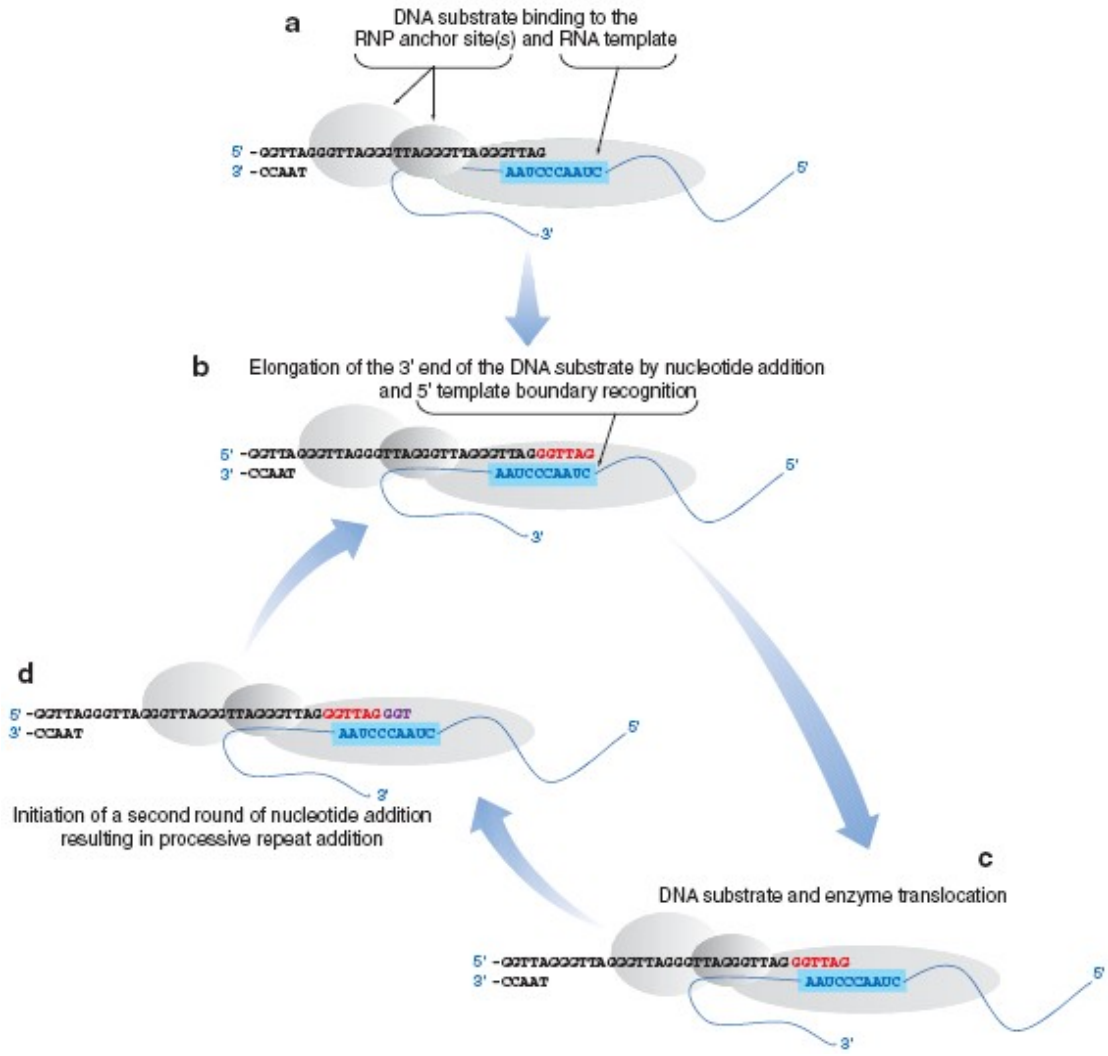

Figure 3
Basic steps of telomere elongation by telomerase
The telomerase ribonucleoprotein is comprised of at least two components: the tert protein depicted in grey and the RNA molecule terc depicted in blue. The basic steps of telomeric DNA elongation include: a) binding of the 3' end of the telomeres to telomerase b) addition of the nucleotides (red) to the telomeric end c) upon reaching the 5' template boundary recognition the telomerase translocates and repositions the 3' of the DNA d) initiation of the following round known as 'processive repeat addition' which is a unique property of telomerase among the polymerases. Adapted from Autexier and Lue, 2006.

Analysis of tert from various species reveals conserved structure of the protein that can be divided into four main functional domains: a tert essential N-terminal domain (TEN), an RNA binding domain, a Reverse Transcriptase domain and a C-terminal domain (Figure 4).
Parts of the TEN domain have been previously referred to as the GQ motif, dissociates activities of telomerase (DAT) domain, RNA interaction domain-1 (RID1), Region I and N-term. The TEN domain high resolution structure was the first portion of tert to become known. The data are obtained from the crystal structure of

Tetrahymena thermophila tert (Jacobs et al. 2006). Groups of residues from the TEN domain have been reported to contribute to repeat addition processivity of telomerase - the property of telomerase to add successive six bps repeats to the substrate without termination or dissociation. Some studies indicate that in this region might be located the template – proximal anchor site where the DNA binds and this way contributing to the repeat addition processivity. In addition this domain has been shown to mediate an interaction with the RNA component of telomerase and more specifically in humans with the pseudoknot template of terc. Mutations in this domain were found to affect the template usage. Part of the TEN domain has been shown to physically interact in vitro with the RT and the C-terminal domain while complementation studies suggest that the same part of the tert molecule can restore activity of another truncated tert molecule that is missing this part (head to tail interaction). These experiments imply a role of the TEN domain in tert multimerization, as well.

The RNA binding domain encompasses three conserved motifs CP, QPF and T identified on the bases of sequence alignment. Motifs CP and T are known to bind the double- and single-stranded RNA regions of the template boundary element, respectively. While this domain is a common RNA binding structure in terts from various species it seems that it interacts with different targets of RNA in different organisms (Autexier and Lue, 2006). In humans the target site appears to be a conserved CR4-CR5 region distant from the template and essential for telomerase activity in vitro (Chen and Greider, 2003).

The Reverse Transriptase (RT) domain includes seven motifs (1, 2, A, B', C, D and E) highly conserved functionally and in sequence. The chrystallographic structures of the RT domain of HIV-1 are available and in comparison with the structure of the T. castaneum tert (Gillis et al. 2008) give information for the human tert RT domain and active site. The RT domain is essential for telomerase activity. The motifs 1, 2, B', C and E have been shown to be implicated in nucleotide addition processivity. Motifs 1 and 2 make contact with the nucleotide triphosphate and motif E has been shown to interact with the primer terminus, also known as the 'primer grip'. Residues essential for metal binding and polymerase chemistry are located within the C motif (Autexier and Lue, 2006). The tert active site is formed by three aspartic acids (Asp) located in motifs A and C. All polymerases are thought to use a two metal mechanism for catalysis and principal to this mechanism are the metal –binding Asp residues near the active site. The similarity observed between these proteins suggests that telomerase uses a two metal mechanism to mediate the nucleotide transfer. The telomerase nucleotide-binding pocket consists of conserved residues that form the motifs 1, 2, A, C, B' and D and are implicated in template and nucleotide binding.

The C-terminus is a weakly conserved part of the protein. It encompasses four regions named E-I, E-II, E-III, and E-IV that were found to be required for in vitro telomerase activity. These regions are residing in the most conserved areas of the C- terminal domain. Region E-I and the upstream motif E of the RT domain form a continuous stretch essential for telomerase activity. In addition, the E-I region has a putative Nuclear Export Signal (NES) and possibly has a role in protein localization. Within the E-II region a 14-3-3 site has been mapped suggesting a role in the activity of the enzyme (Banik et al. 2002).

The RNA components of telomerase are very different in sequence and in length among species. The size ranges from small molecules (148-209nts for ciliates) to larger ones (930–1544nts for yeast). The human terc consists of 451nts and have a conserved secondary structure. The secondary structure of vertebrate RNA consists of four highly conserved structural domains: the pseudoknot, CR4–CR5, box H/ACA,

and CR7 domains. The pseudoknot and CR4–CR5 domains can independently assemble with tert protein and are essential for telomerase activity in vitro. The box H/ACA domain is important for RNA processing and RNA stability in vivo. The base-paired structure of helix P1b is essential for correct utilization of the template (Figure 5) (Chen and Greider, 2003).

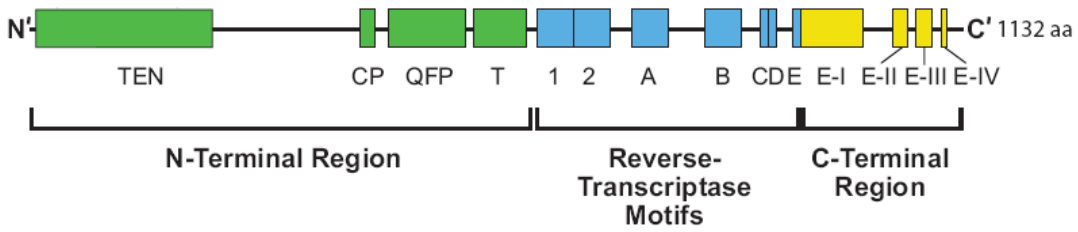

Figure 4
The functional structure of human tert is schematically depicted. The tert protein can be roughly divided into four domains. The TEN domain, the RNA binding domain containing the conserved motifs CP,QFP and T (also known as hypomutable domains II, III and IV ,respectively), the RT domain containing the seven conserved RT motifs and the weakly conserved C-terminal region.

Dyskerin, a nucleolar protein encoded by the *DKC1* gene on the X chromosome, is present in small nucleolar ribonucleoprotein particles and modifies specific uridine residues of ribosomal RNA by converting them to pseudouridine. Dyskerin is functionally associated with the RNA component of telomerase. It was shown that it is required for proper folding and stability of terc (Dokal et al. 2000) and to be part of the basic human telomerase enzyme complex (Cohen et al. 2007).

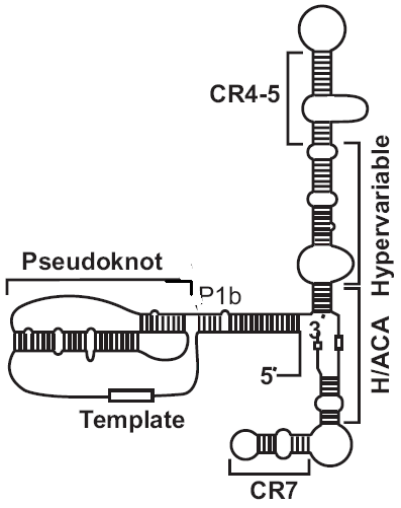

Figure 5
The secondary structure of terc consists of four highly conserved structural domains: the pseudoknot, CR4-CR5, boxH/ACA and CR7 domain.

Telomerase has an essential role in maintaining telomere length as it is the main mechanism for telomere elongation. There are increasing evidence in the literature

that telomerase may have other functions independent of the telomere maintenance involving apoptosis, DNA repair, gene expression regulation and stem cell function (Cong and Shay. 2008).
A telomerase deficient animal model was generated to better understand the functions of telomerase in vivo. That is a knock out transgenic mouse for the murine *TERC* gene. The mice are viable but show progressive telomere loss with each successive generation starting as soon as the third generation. These mice show compromised ability to respond to physiological stress and this lead to some phenotypes that can be linked to the telomere shortening. These phenotypes include hair graying, alopecia and skin lesions, male and female infertility, embryonic mortality due to a defective closure of the neural tube, spleen atrophy and reduced proliferation of B and T lymphocytes, small size and severe intestinal atrophy, reduced proliferative potential of the bone marrow stem cells, heart dysfunction, reduced proliferative capacity of adult neural stem cells, reduced longevity due to spontaneous tumor formation and impaired wound healing. These mice did not develop generalized premature aging as they have no signs of osteoporosis, cataract formation, atherosclerosis or impaired glucose tolerance (Blasco, M. A. 2005). The features that are associated with the short telomeres can be prevented if telomerase is re-introduced to these *TERC*-/- mice and elongates the short telomeres. In addition it was proposed that it is the individual critically short telomeres that are responsible for the phenotype rather than the average telomere length (Hemann et al.2001). Although the results from mice cannot be extrapolated directly to human as murine telomere length is much longer and the telomerase pattern of expression in mice tissues is wider, the short telomere mouse model can speculate that telomerase deficiency in human may contribute to aging phenotypes as well.
By the time we started the present study there were experimental evidence that telomere shortening is the cause for two human disorders: dyskeratosis congenita and bone marrow failure.
Dyskeratosis congenita (DKC) is an inherited disease characterized by a triad of abnormal skin pigmentation, nail dystrophy and leucoplakia. It is a rare disorder occurring in 1 out of 1000000 individuals in the general population. DKC is a multisystem syndrome characterized by haematological, pulmonary, gastrointestinal, endocrine, skeletal, neurological, urologic and immunologic abnormalities. Bone marrow failure appears to be the principal cause of death as well as pulmonary complications and a predisposition to malignancy while the median age of death is 16years. In general the abnormalities are not present at birth but develop later in life. There is significant variability in the severity and the clinical phenotype of the syndrome manifestations among patients even within the same family. Although majority of the affected are males suggesting an X-linked pattern of inheritance, there are families that exhibit dominant or recessive autosomal form of the disease (Dokal et al. 2000). Using linkage analysis in a large family the gene responsible for the X-linked form found to be *DKC1* and mapped to the Xq28 (Heiss et al. 1998). *DKC1* encodes for dyskerin the protein that is part of the basic human telomerase complex. Most of the *DKC1* mutations found in DKC affected are missence mutations predicted to have an impact in binding of dyskerin to the RNA component of telomerase. In addition to *DKC1*, mutations in *TERC* (Vulliamy et al. 2001) and *TERT* (Armanios et al. 2005) have been found in patients with DKC. In both cases the families had autosomal dominant pattern of disease transmission with genetic anticipation for some phenotypes. A missence mutation of *NOLA3*, a gene encoding for the protein nop10 that is associated with the telomerase complex, has been identified as the cause in a

family with autosomal recessive DKC (Walne et al. 2007). Mutations in two more proteins associated with the telosome nhp2 (Vulliamy et al. 2008) and tinf2 (Walne et al. 2008) have been detected in patients with DKC. Patients with DKC have very short telomeres independently of the type of mutation or the pattern of inheritance indicating that there is a common mechanism involving telomere shortening. Overall <50% of all DKC patients seem to have a mutation in one of the above genes involved in telomere maintenance (Garcia et al. 2007).

DKC can be considered as a bone marrow failure inherited syndrome as 80% of the patients develop aplastic anemia. Sequencing *TERC* in patients with aplastic anemia that did not have any manifestations of DKC led to the identification of different mutations. Furthermore these patients had very short telomeres (Vulliamy et al. 2002) indicating that the bases to the disorder may be telomere maintenance. Various *TERC* mutations have been found in various cases of bone marrow failure syndromes such as constitutional aplastic anemia, paroxysmal nocturnal hemoglobinuria, and myelodysplastic syndrome. Screening other components of telomerase complex in patients with aplastic anemia led to identification of *TERT* mutations (Yamaguchi et al. 2005). Recently mutations in *TERT* were found in patients with acute myeloid leukaemia, as well (Calado et al. 2009).

Herein, with the work presenting in the following pages we provide experimental evidence associating another disorder, adult onset idiopathic pulmonary fibrosis, with telomere shortening and telomere maintenance.

MATERIALS AND METHODS

Clinical studies

We have collected 46 families with at least two members affected with interstitial lung disease.34 out of the 46 families had individuals with IPF. In addition, 44 sporadic cases with interstitial lung disease, 31 of which met the criteria for IPF (ATS/ERS, 2002), were included in the study. All sporadic cases had no affected first or second degree family members. A medical questionnaire was completed by each participant. Medical records including pulmonary function tests, chest x-ray, computed tomography of the chest, high resolution computed tomography of the chest, transbronchial biopsy of the lung and open lung biopsy as well as blood samples were obtained when available. All participants provided written informed consent according to study protocol approved by the University of Texas Southwestern Institutional Review Board.

Genomic DNA extraction

Genomic DNA was purified from white blood cells using an Autopure LS (Qiagen,Valencia,CA). In four cases DNA was extracted and amplified from formalin fixed paraffin embedded tissue. Tissue shavings were deparaffinized with 1ml of xylene and then dehydrated with 1ml xylene: ethyl ethanol and finally 1ml ethyl ethanol. DNA was isolated from the dried pellets using the tissue protocol of the QIAamp DNA Mini Kit (Qiagen) and amplified according to the GenomePlex Whole Genome Amplification Kit (Sigma, St. Louis, MO).

Whole genome SNP linkage scan

Whole genome SNP linkage scan was performed on Families 11 and 31. DNA extracted and amplified from biopsied lung tissue or DNA from whole blood was genotyped for over 6000 SNP markers for each individual. Individuals with IPF, pulmonary fibrosis, or unclassified pulmonary disease were considered ‘‘affected,’’ and all others were assigned an unknown affectation status. The genotyping was performed by the University of Texas Southwestern Microarray Core using the Illumina Linkage IVb SNP panel. The panel contains more than 6,000 polymorphic SNPs distributed evenly through the whole human genome. The percentage of total calls to the possible total calls – call rates- varied from 99.7–100% for Autopure-purified DNA and between 72–98% for whole-genome-amplified DNA extracted from archival samples. Multipoint Engine for Rapid Likelihood Inference (MERLIN) (Abecasis et al. 2002) is the computer program that was used to screen the genome and perform haplotyping. Linkage analysis was done based on a model-free method (Kong et al. 1997) followed by evaluation of the regions with the highest signals by using a model of autosomal dominant pattern of inheritance method. The LOD score was 2.82 ($P<0.0002$) model-free method and 2.68 ($P<0.002$) model –based method. The 1-LOD drop interval extended from the end of chromosome 5p15 to SNP rs959937. That is a region of 4.3Mb containing 45 genes. The LOD score 2.82 ($P<0.0002$) model-free method and 2.68 ($P<0.002$) model –based method.

Sequencing and mutation analysis

Polymerase chain reaction PCR amplification of *TERT* and *TERC* was performed using genomic DNA as template. The primers were designed to sequence all 16 exons and splice sites of *TERT* using ELXR (Exon Locator and Extractor for Resequencing) (Schageman et al. 2004). Once PCR product was amplified the remaining dNTPs and primers were removed from the mixture with recombinant exonuclease I and shrimp alkaline phosphatase (Exo-Sap, USB) digestion. The purified PCR product was sequenced on an ABI 3700 automated sequencer by BigDye terminator cycle sequencing reagents (Applied Biosystems). All sequences were read both directions using a forward and a reverse primer and mutations were confirmed in three separately amplified PCR products. *TERC* and *TERT* were sequenced in their entity for all probands and then for probands' family members when a mutation was detected. All sequences used in the comparative alignment were obtained from the NCBI web site at www.ncbi.nlm.nih.gov.
The primer pair for *TERC* gene was forward 5'-tcatggccggaaatggaact-3' and reverse 5'-gggtgacggatgcgcacgat-3'and annealing temperature 60°C. The amplification of *TERC* was performed in the presence of Advantage –GC Genomic polymerase (Clontech), 1.0M GC-Melt reagent and 1.1mM $MgCl_2$ and the final size of the product was 653bps. PCR conditions and primers for amplifying the *TERT* gene from genomic DNA are listed in the following table.

TERT exon	Forward primer	Reverse primer	Amplicon size	Taq polymerase	[Mg^{+2}]	Tm Annea
1	5'-agcccctccccttccttt-3'	5'-ctccttcaggcaggacacct-3'	451	Advantage GC	1.1	60
2a	5'-accagcgacatgcggaga-3'	5'-GTCGCCTGAGGAGTAGAGGAA-3'	844	Advantage GC	1.1	60
2b	5'-GGTGTACGCCGAGACCAAG-3'	5'-CGTTCGTTGTGCCTGGAG-3'	486	Advantage GC	1.1	58
2c	5'-CCGAGGAGGAGGACACAGAC-3'	5'-aaaccgcgtgtccatcaa-3'	493	Ampplitaq Gold	1.5	60
3	5'-ccttggtgagctggatgtg-3'	5'-cacagaatccacttggaccag-3'	455	Ampplitaq Gold	1.5	55
4	5'-gtgagcttcccctagtctgt-3'	5'-ataccaaatgtggggctcaa-3'	307	Ampplitaq Gold	1.5	55
5	5'-tttcaaacagggtctgaggaa-3'	5'-tgtgtcctcaacagtgacagg-3'	478	Ampplitaq Gold	1.5	55
6	5'-ggcagaggtgatgtctgagtt-3'	5'-agatacatgcaccacgacaca-3'	385	Ampplitaq Gold	1.5	55
7	5'-ggagtcccaggtgtgtctgta-3'	5'-caaggcacacagctcatcat-3'	454	Ampplitaq Gold	1.5	55
8	5'-cgcacttcatcacaaacactg-3'	5'-ccagaaaaggagactctggtg-3'	328	Ampplitaq Gold	1.5	55
9	5'-gctgaatggtagacgtgtcgt-3'	5'-cactgaatgcatcaaaagcaa-3'	299	Ampplitaq Gold	1.5	55
10	5'-agaattgcacaagctgatggt-3'	5'-gagaggacttggcagagacaa-3'	419	Ampplitaq Gold	1.5	55
10	5'-agtccctcgtagacagatacta-3'	5'-gcctcacCTGAGGAAGGTTT-3'	149	Ampplitaq Gold	1.5	55
11	5'-tgctccaaatcaccacttctc-3'	5'-cctcactcccacagaaagatg-3'	486	Ampplitaq Gold	1.5	55
12	5'-gatggcatgtagcatttggag-3'	5'-ccatgccttcatgtacacaca-3'	486	Ampplitaq Gold	1.5	55
13	5'-tctcctggttccttcctgtct-3'	5'-agacattccttgcccctaaaa-3'	422	Ampplitaq Gold	1.5	55
14	5'-tgcgtgttcatacagatggtg-3'	5'-tcctaagcccagattcactca-3'	461	Ampplitaq Gold	1.5	55
15	5'-ggaaatttcacctggagaagc-3'	5'-ccagcgtttaatcacatagggg-3'	481	Ampplitaq Gold	1.5	55
16	5'-gtcctaggagggttggaggat-3'	5'-ATTCCTATGTGGGGAGTGGAA-3'	489	Ampplitaq Gold	1.5	55

Lowercase letters indicate intron sequences while exon sequences are indicated by the uppercase letters. Advantage –GC Genomic polymerase (Clontech). Amplitaq Gold (Applied Biosystems).The 2a amplicon was sequenced using the following internal primers: 5'-ACGCTAGTGGACCCCGAAG-3' and 5'-CCCTGACGCTATGGTTCCAG-3'.

Mutagenesis of *TERT* and *TERC*

The mutations c.97C>T, c.2593C>T, c.2594G>A, c.430G>A, c.1456C>T and c.3346_3522del were introduced into the parental plasmids pGRN121 (the ecoRI fragment from lamda clone 25-1.1.6 containing the entire cDNA encoding human tert protein was inserted into the ecoRI site of pbluescriptIISK+ such that the 5' end of the cDNA is near the T7 promoter in the vector) and pGRN125 (the tert coding sequence was inserted into the Not1 site of pbbs235 so that the open reading frame is in the opposite orientation than the Lac promoter) using the QuikChange site-directed mutagenesis kit (Stratagene).
The c.2240delT mutation was generated by PCR, digested with NdeI , EcoRI and then directionally subcloned into pCITE-4a (Novagen). The PCR reaction consisted of three successive steps. We used pGRN125 as template for the first and second reaction. Primer pairs were forward1a: 5'-aagcttcatatgccgcgctccccgct -3'and reverse1a: 5'-atgggcggccttctggccacggcataccgacgca-3' for the first reaction and forward1b: 5'-tgcgtcggtatgccgtggccagaaggccgcccat-3' and reverse 1b: 5'-ggaattcttagtccaggatggtcttgaagtc-3' for the second one. Then the PCR products from the first two steps served as template for the third reaction with the same set of primers as the first one (forward1a, reverse1a). Wild type control was generated with PCR and subcloned in the same restriction sites.
TERC was amplified from genomic DNA by PCR using oligonucleotides 5'-ggggaagctttaatacgactcactatagggttgcggagggtgggcctg-3' and 5'-ccccggatcctgcgcatgtgtgagccgagtcctggg- 3', digested with HindIII and BamHI, and subcloned into pUC18 to generate plasmid pKT26. *TERC* r.37a>g was introduced in pKT26 with the use of the QuikChange site-directed mutagenesis kit (Stratagene). The sequence in all clones was verified by sequencing.

In vitro reconstitution of Telomerase

Human telomerase complexes were reconstituted in vitro by using the TnT transcription/
translation system (Promega, Madison, WI). Briefly, linearized TERT constructs (0.5 μg) and FspI-linearized telomerase RNA constructs (0.25 μg) were used as substrates and co-expressed in 25μl TnT reactions according to the manufacture's instructions.

TRAP Assays

Telomerase activity of the in vitro reconstituted complex was assayed with the amplification TRAP protocol as described (Herbert t al. 2006, Weinrich et al. 1997). The TnT reactions were serially diluted 1:5 in 1xTRAP buffer (200 mM Tris·Cl, pH 8.3, 15 mM MgCl2, 630 mM KCl, 0.5% (v/v) Tween 20, 10 mM EGTA) starting with 0.3–0.5 μl of the 25μl reaction. Then the substrate (diluted TnT reaction) was mixed with Trap 10x buffer, Primer Mix (100ng), TS CY5 (100ng), dNTPs 0.025mM, Ultra pure BSA (0.4mg/ml), Flexi taq (promega) (2u/50μl) and RNase free water. The primer mix includes the reverse primer (ACX = 5′- gcgcggcttacccttacccttacccctaacc -3′), the substrate for the 36-bp internal standard control (TSNT = 5′-aatccgtcgagcagagttaaaaggccgagaagcgat -3′), and the reverse primer for the internal standard (NT = 5′- atcgcttctcggcctttt -3′). The TS CY5 sequence is 5'-Cy5-aatccgtcgtcgagcagagtt. One wild type sample of the highest concentration was first heat inactivated at 85°C for 10 min and then incubated at room temperature for 30min

with the rest of the reactions. This sample served as the negative control. The positive control was 250 cell equivalents of H1299, human lung carcinoma cells known to be positive for telomerase. Subsequently each sample was treated with 1 unit of RNaseH (Invitrogen) for 45 min at 37°C to degrade all the RNA products from the TnT reaction before proceeding to PCR amplification of the extension products. Half amount of each reaction was loaded in a 10% non denaturing acrylamide gel and run at 20V cm^{-1}.The gel was scanned with a Typhoon Phosphorimager that can read the Cy5 fluorescence and the intensity of the 6-bp incremental ladder and the 36-bp internal control were determined with Image Quant software. Relative amounts of telomerase activity were calculated as the ratio of the intensity of the sample's telomerase activity to that of the internal control band and normalized to the positive control of the experiment.

Mixing experiment

V747fs *TERT* and wt *TERT* were combined in the same TnT reactions at the following ratios 1:0, 0.75:0.25, 0.5:0.5, 0.25:0.75, 0:1 while the total amount of *TERT* constract used was 0.5μg. The TRAP reactions were run as described. Parallel TnT reactions were designed using the radioactive label [^{35}S] methionine to confirm the expression of the proteins in the indicated amounts. The reactions were loaded on a 7% SDS/PAGE gel. Once the electrophoresis was completed the gel was fixed in 50% methanol and 10% glacial acid for 1 h, soaked in En^{3}hance Kodak (Perkin Elmer) for an additional hour, vacuum dried and finally exposed to film at -80°C for 4-16hours.

Terminal Restriction Fragment Analysis (TRF)

Genomic DNA isolated from leukocytes was assayed for TRF analysis. Briefly the DNA was digested with a mixture of 6 restriction enzymes and separated in an agarose gel. The gel was denatured and prehybridized before hybridization with the radiolabeled telomeric repeat probe. Then washed with SSC prior to exposure to a phosphor screen and scanned with a PhosphorImager. ImageQuant software was used to visualize the probes and to calculate the mean telomere length as already described (Herbert et al. 2003) with a slight modification. We placed a grid of 200boxes, instead of 30, over each lane and then determine the signal intensity and the size (according to the DNA marker) corresponding to each box. The percentage of short telomeres was calculated by dividing the relative signal intensity of each lane (arbritrarily defining the short telomeres between 1.9 and 4.3 kb) by the relative signal intensity of the entire lane (between 1.9 and 19 kb).

RESULTS

Out of the 46 families that we have collected with idiopathic interstitial lung disease we picked two of the largest Caucasian ones to perform whole genome linkage scan: Family 11 and Family 31. The proband of Family11 was a 58year old male that presented with cough and dyspnea. He had crackles on physical examination and the high resolution computerized tomography (HRCT) of the lungs showed peripheral upper and lower lobe interstitial fibrosis typical for IPF. His brother had the same symptoms and HRCT revealed a pattern of peripheral lower lobe predominant interstitial fibrosis. Open lung biopsy was positive for usual interstitial pneumonia (UIP) confirming the diagnosis of IPF. The patient underwent lung transplantation. Their father died from respiratory failure, bronchiectasis and pulmonary fibrosis at the age of 56 (information obtained from death certificate). Their aunt reported cough and progressive dyspnea for one year with rapid worsening the last four weeks before her death at 66year old. The computerized tomography of the lungs showed peripheral lower lobe pulmonary fibrosis with superimposed ground glass opacity while the open lung biopsy had features of chronic interstitial pneumonia. The patient did not respond to immunosuppression therapy. Their other aunt suffered from sarcoidosis with pulmonary, dermatologic and neurologic manifestations and her lung disease appeared to be responding to steroids and immunosuppression. The proband's grandfather presented with cough and dyspnea and a suspected atypical mycobacterial infection. The chest radiography had features of pulmonary fibrosis while his death certificate showed pneumonia, pneumothorax, emphysema and honeycomb lungs (terminal lung disease). The brother of the grand father died at 62 from acute respiratory failure, severe hypoxemia and chronic obstructive pulmonary disease (COPD). One sister of the grandfather died at 72. She had pulmonary fibrosis with lung biopsy positive for UIP, osteoporosis with multiple vertebral fractures and anemia. The other sister who died at 68year old was a non smoker with longstanding dyspnea by family report. Her daughter died when she was 55year old with worsening dyspnea for the last two years. Nine years prior her death she was diagnosed with breast cancer. She went through left side mastectomy and therapy with cytoxan, methotrexate, and 5FU. HRCT showed extensive peripheral interstitial changes and honeycombing while the lung biopsy had features of UIP. She did not respond to immunosuppression treatment in agreement with the IPF diagnosis. Spirometry of the lung was performed on the proband, his brother and their aunt all showing reduced diffusing capacity of the lung for carbon monoxide.
The proband of the Family 31 was a 67year old male smoker who presented with dyspnea and crackles. HCRT pictures showed interstitial fibrosis predominately in the periphery of lower lobes consistent with IPF. His cousin was a nonsmoker who showed progression of pulmonary fibrosis mainly in the upper lobes over a period of four years by chest x-ray. The transbronchial lung biopsy was not diagnostic and she suffered from vertebral compression fracture by osteoporosis and anemia. The other cousin showed similar phenotype with progression of pulmonary fibrosis favoring the upper lobes and multiple vertebral compression fractures. His father died at the age of 59 and was a smoker diagnosed with emphysema by family report. One of proband's uncles died by respiratory failure while the other one by pneumonia at the age of 34 both information obtained by family report. His aunt had a pulmonary disease "suspected tuberculosis" and died at 34 years old in an accident (supplement Tsakiri et al. 2007).

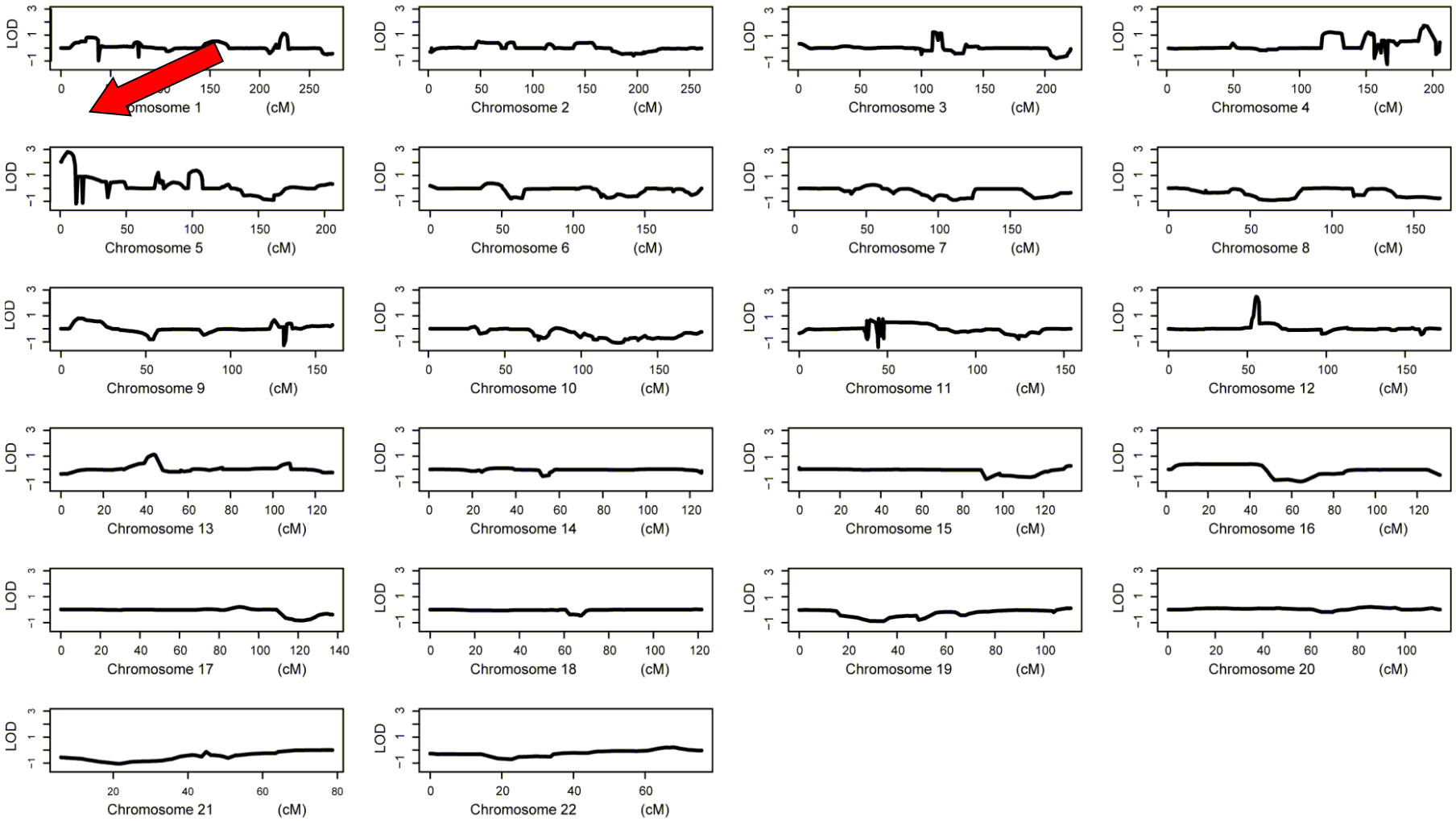
LOD
Chromosome 1 (cM)
Chromosome 2 (cM)
Chromosome 3 (cM)
Chromosome 4 (cM)
Chromosome 5 (cM)
Chromosome 6 (cM)
Chromosome 7 (cM)
Chromosome 8 (cM)
Chromosome 9 (cM)
Chromosome 10 (cM)
Chromosome 11 (cM)
Chromosome 12 (cM)
Chromosome 13 (cM)
Chromosome 14 (cM)
Chromosome 15 (cM)
Chromosome 16 (cM)
Chromosome 17 (cM)
Chromosome 18 (cM)
Chromosome 19 (cM)
Chromosome 20 (cM)
Chromosome 21 (cM)
Chromosome 22 (cM)

Figure 6
Whole genome linkage scan for F11 and F31 results are presented for each chromosome. The LOD scores are plotted against chromosome distance expressed as units of recombinant frequency (cM). The red arrow is pointing the peak at the end of chromosome 5.

Genotyping over 6000 markers for each individual of Family 31 and Family 11 and analysis of the haplotypes result in regions that might show genetic linkage. The highest peak that did not quite reach statistical significance was at the end of chromosome 5 (Figure 6). The logarithm of odds (lod) score was 2.82 (model free) and 2.68 (using a model of autosomal pattern of inheritance) (Figure 7). Within the1-lod drop interval there are 45 genes (Table 3).

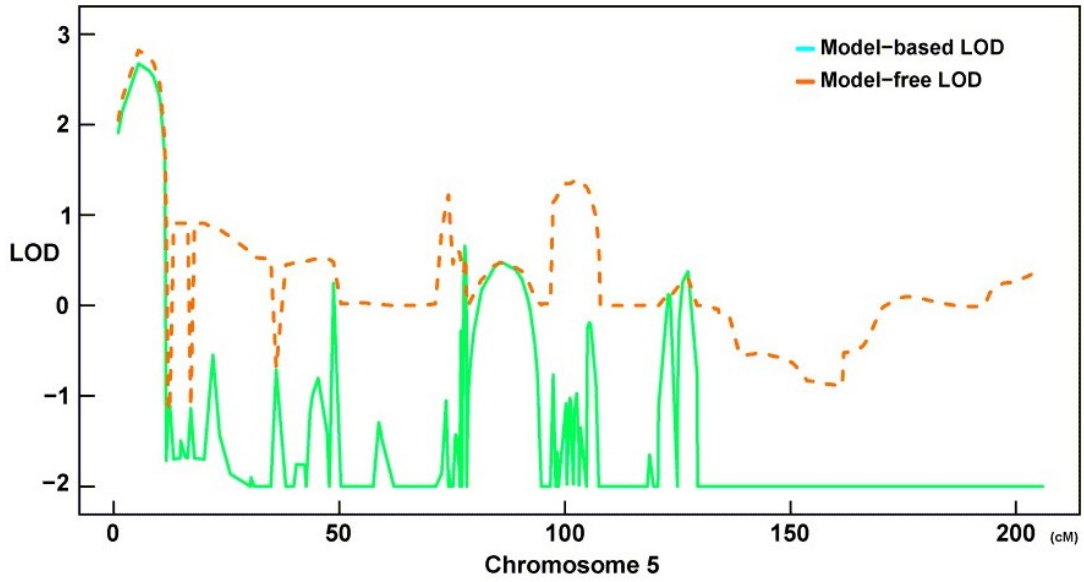

Figure 7
The model free estimation of LOD gives a slightly higher score at the end of chromosome 5 compare to the autosomal dominant pattern of inheritance based model.

Table 3

	Gene Symbol	Description
1	*IRX1*	iroquois homeobox 1
2	*LOC100132531*	hypothetical LOC100132531
3	*LOC100130748*	hypothetical LOC100130748
4	*C5orf38*	chromosome 5 open reading frame 38
5	*IRX2*	iroquois homeobox 2
6	*LOC100133292*	hypothetical protein LOC100133292
7	*LOC731559*	hypothetical protein LOC731559
8	*IRX4*	iroquois homeobox 4
9	*NDUFS6*	NADH dehydrogenase (ubiquinone) Fe-S protein 6, 13kDa (NADH-coenzyme Q reductase)
10	*MRPL36*	mitochondrial ribosomal protein L36
11	*LOC100132773*	hypothetical protein LOC100132773
12	*LOC728613*	programmed cell death 6 pseudogene
13	*LOC100128216*	hypothetical protein LOC100128216
14	*LPCAT1*	lysophosphatidylcholine acyltransferase 1
15	*SLC6A3*	solute carrier family 6 (neurotransmitter transporter, dopamine), member 3
16	*CLPTM1L*	CLPTM1-like
17	*TERT*	telomerase reverse transcriptase
18	*SLC6A18*	solute carrier family 6, member 18
19	*LOC100128907*	hypothetical protein LOC100128907
20	*SLC6A19*	solute carrier family 6, member 19
21	*SLC12A7*	solute carrier family 12 (potassium/chloride transporters), member 7
22	*NKD2*	naked cuticle homolog 2 (Drosophila)
23	*TRIP13*	thyroid hormone receptor interactor 13
24	*BRD9*	bromodomain containing 9
25	*LOC100132536*	hypothetical LOC100132536
26	*LOC100129204*	hypothetical LOC100129204
27	*ZDHHC11*	zinc finger, DHHC-type containing 11
28	*ZDHHC11B*	zinc finger, DHHC-type containing 11B
29	*LOC100132605*	hypothetical LOC100132605
30	*TPPP*	tubulin polymerization promoting protein
31	*CEP72*	centrosomal protein 72kDa
32	*SLC9A3*	solute carrier family 9 (sodium/hydrogen exchanger), member 3
33	*LOC25845*	hypothetical LOC25845
34	*EXOC3*	exocyst complex component 3
35	*LOC116349*	hypothetical protein BC014011
36	*LOC100129491*	hypothetical protein LOC100129491
37	*AHRR*	aryl-hydrocarbon receptor repressor
38	*LOC100129064*	hypothetical LOC100129064
39	*PDCD6*	programmed cell death 6
40	*SDHA*	succinate dehydrogenase complex, subunit A, flavoprotein (Fp)
41	*CCDC127*	coiled-coil domain containing 127
42	*FLJ44896*	FLJ44896 protein
43	*LOC389257*	leucine-rich repeat-containing protein 14-like
44	*PLEKHG4B*	pleckstrin homology domain containing, family G (with RhoGef domain) member 4B
45	*LOC100128803*	hypothetical LOC100128803

We focused on one of these genes, *TERT*, that seemed to be an excellent candidate for pulmonary fibrosis. A null mutation in a highly conserved domain of tert was found to be associated with autosomal dominant form of dyskeratosis congenita (DKC) in a three generation pedigree (Armanios et al. 2005). DKC is a rare inherited disorder characterized by a triad of skin hyperpigmentation, nail dystrophy and mucosal leukoplakia. It is a multisystem disease in which 20% of the patients present with idiopathic pulmonary fibrosis. Pulmonary complications are the second cause of death for these patients following bone marrow failure (Dokal et al.2000). Four out of seven affected members of this family displayed pulmonary fibrosis as a manifestation of the disease. The authors remark that the affected individuals showed idiopathic pulmonary fibrosis and that this disorder was characterized by anticipation. In addition pulmonary fibrosis was the only manifestation of the disease in one carrier of *TERT* mutation (Armanios et al. 2005). These observations justify *TERT* as an excellent gene for IPF located in the linkage window.

Next, we sequenced the 16 exons and the consensus splicing sites of the *TERT* for the probands of Families 11 and 31. The proband of Family 11 was heterozygous for a transition mutation (c*g*t to c*a*t) in codon 865. This mutation is predicted to change a highly conserved arginine to histidine which is part of the consensus sequence of motif C. Motif C is one of the seven reverse transcriptase (RT) motifs that are essential for telomeric DNA synthesis. RT motifs are conserved from yeast to human (Nakamura et al. 1997). We identified a deletion of thymidine at position 2240 in the cDNA of the proband of Family 31. This frameshift mutation disrupts the reading frame and is predicted to result in a truncated protein missing half the RT domain and the entire C- terminal region. Then we sequenced the rest of the family members. In both families we found that the mutations and the phenotype of the disease were segregating together. Characteristically, the aunt of the proband of Family 11 that had pulmonary disease as a manifestation of sarcoidosis was negative for the mutation. *TERT* was then sequenced in the other 44 probands of our families' collection. Four more new mutations were identified. Specifically, the proband of Family 34 was found heterozygous for a missense mutation (*c*gc to *t*gc) in codon 486. This mutation is predicted to change a highly conserved arginine to cysteine within the conserved QFP domain of the N-terminal region of tert. In addition to that the proband of Family 40 was heterozygous for a missence mutation (*c*cc to *t*cc) in codon 33 resulting in a serine instead of a proline in the telomerase essential N- terminal domain (NET) of tert. In addition to P33S another missense mutation was detected in the NET domain of tert. The proband of Family 71 was heterozygous for (*g*tg to *a*tg) in nucleotide 430 that results in methionine instead of valine in codon 144. V144M is adjacent to a completely conserved glycine (Gly145) suggesting steric effects. A 177bps deletion was found in the proband of Family 8. This deletion disrupts the reading frame of the protein and results in a protein missing the last 16aa and extending in the 3' untranslated region of *TERT* (Figure 8).

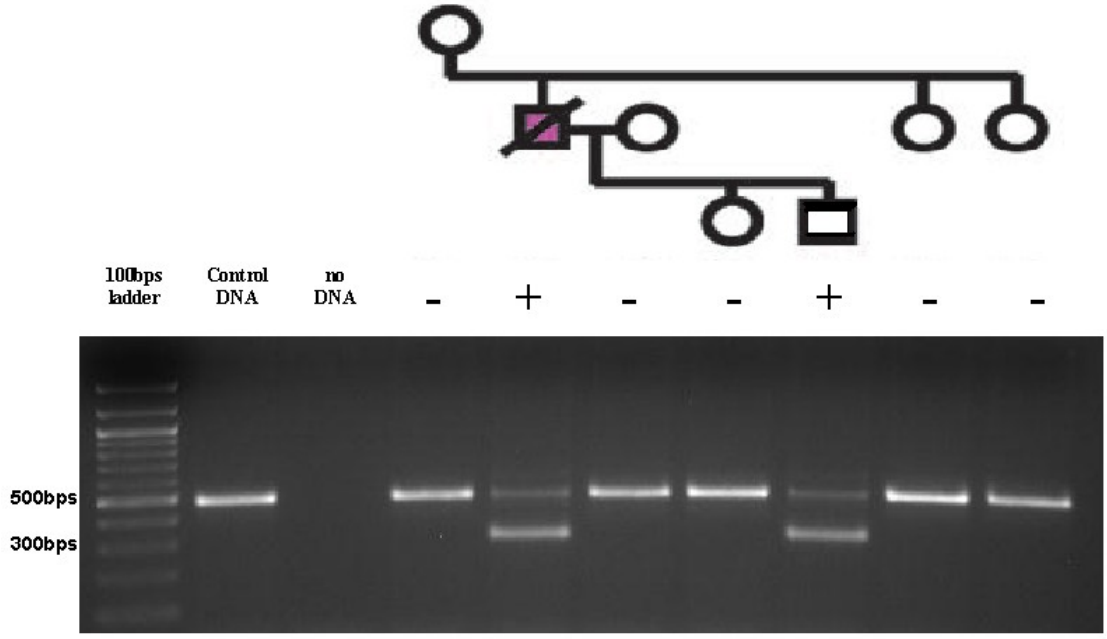

Figure8
c.3346_3522del mutation in Family 8. Oligos spanning exon 16 of TERT are designed to give an amplicon size of 489bps. The proband and his father are both heterozygous for the 177bps deletion. The PCR product from genomic DNA is run in an agarose gel. The product that is missing 177bps is migrating faster in the gel resulting in a double band for the carriers. The presence (+) or absence (-) of E1116fsX1127 is indicated above each lane.

Then we sequenced 44 individuals with interstitial lung disease that did not have any affected first or second degree family members. One sporadic case was found to be heterozygous for a missense mutation (*c*gt to *t*gt) that is predicted to change an arginine to cysteine R865C. This mutation involves the same highly conserved arginine in motif C of the reverse trascriptase motifs as the one detected in Family 11 R865H mutation.
To summarize we have identified 6 *TERT* mutations in a total of 46probands (Figure 9). Thus in approximately 13% of familial cases the disease is caused by mutations in *TERT*. An additional *TERT* mutation was detected in a sporadic case out of 44 individual patients screened. Our control group was composed of 94 locally collected ethnically matched individuals, 200 patients with aplastic anemia and a multiethnic panel of 528 healthy individuals as described (Yamaguchi et al. 2005). Either the patients with familial or sporadic cases have the characteristic triad of mucocutaneous lesions seen in DKC or have bone marrow failure. This way the phenotype- interstitial lung disease- of our group is distinguished from DKC or aplastic anemia where *TERT* mutations have been previously described (Armanios et al. 2005, Yamaguchi et al. 2005).
Next we wanted to investigate whether mutations in the RNA component of telomerase *TERC* occur in patients with idiopathic pulmonary fibrosis. *TERC* mutations have been identified in autosomal dominant DKC (Vulliamy et al. 2001) the hereditary disorder that is related to premature death from aplastic anemia or

pulmonary fibrosis. *TERC* mutations were also reported in cases of aplastic anemia (Vulliamy et al. 2002).
In order to do so we sequenced *TERC* in the same group of 46 family probands and 44 sporadic cases. The proband of Family 61 was heterozygous for an adenosine to gouanine substitution at position 37. This arginine is mapped on the terminal residue of P1b helix located upstream of the template. Helix P1b is important for template boundary definition in telomerase and disruption of its sequence results in misincorporation of inappropriate nucleotides (Chen et al. 2003) (Figure 9). This mutation has been previously reported in an individual with DKC and severe bone marrow failure who was a compound heterozygote for variants in *TERC* (Ly et al. 2005). The 94 locally collected individuals were negative for the *TERC* mutation. In families that mutations were detected we sequenced *TERT* or *TERC* for all the family members that genomic DNA was available. There were some carriers that did not have pulmonary fibrosis.

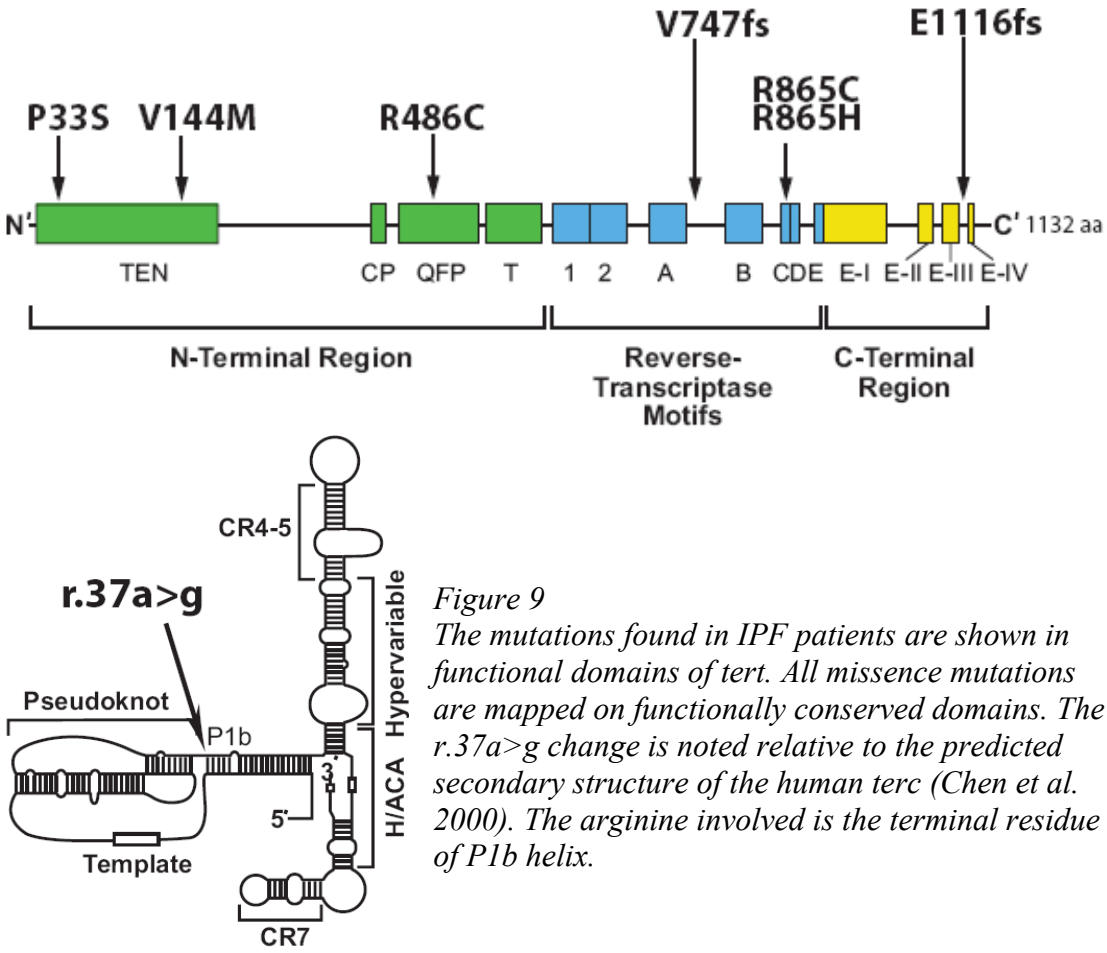

Figure 9
The mutations found in IPF patients are shown in functional domains of tert. All missence mutations are mapped on functionally conserved domains. The r.37a>g change is noted relative to the predicted secondary structure of the human terc (Chen et al. 2000). The arginine involved is the terminal residue of P1b helix.

In order to evaluate the effect of these mutations to the activity of telomerase we used the telomeric repeat amplification protocol (TRAP). That is a PCR based method to detect and measure telomerase activity. It consists mainly of two parts. First the telomerase present in the sample uses the dNTPs to add telomeric repeats to a substrate. The next step involves PCR amplification of these products with a specific designed pair of oligos. The telomerase activity is measured as the intensity of the six bps incremental ladder when the PCR product is run in an acrylamide gel (Herbert et al. 2006). Since tert protein and the telomerase RNA terc are adequate to reconstitute telomerase activity in vitro (Wenrich et al. 1997) we designed experiments

coexpressing recombinant tert and terc and then subject to TRAP assay (Figure 10).We tested all the mutated proteins and found that the activity ranges from almost 0% to 100% compare to the wild type. The two deletion *TERT* mutations showed almost no detectable telomerase activity. The V747fsX mutation gives rise to a protein missing half the reverse transcriptase domain where the catalytic subunit of telomerase is located and as expected the telomerase activity was very low. The E1116fsX mutation is predicted to be missing the C-terminal E-IV domain where the primer grip of the telomerase is mapped (Banik et al. 2002) and the activity in TRAP assay was very little detectable. The five missence mutations had activity 20% to 100%. The *TERC* r.37a>g mutation was almost 75% compare to wild type without a dramatic reduction in telomerase activity as shown before (Ly et al. 2005) (Figure 10,11).

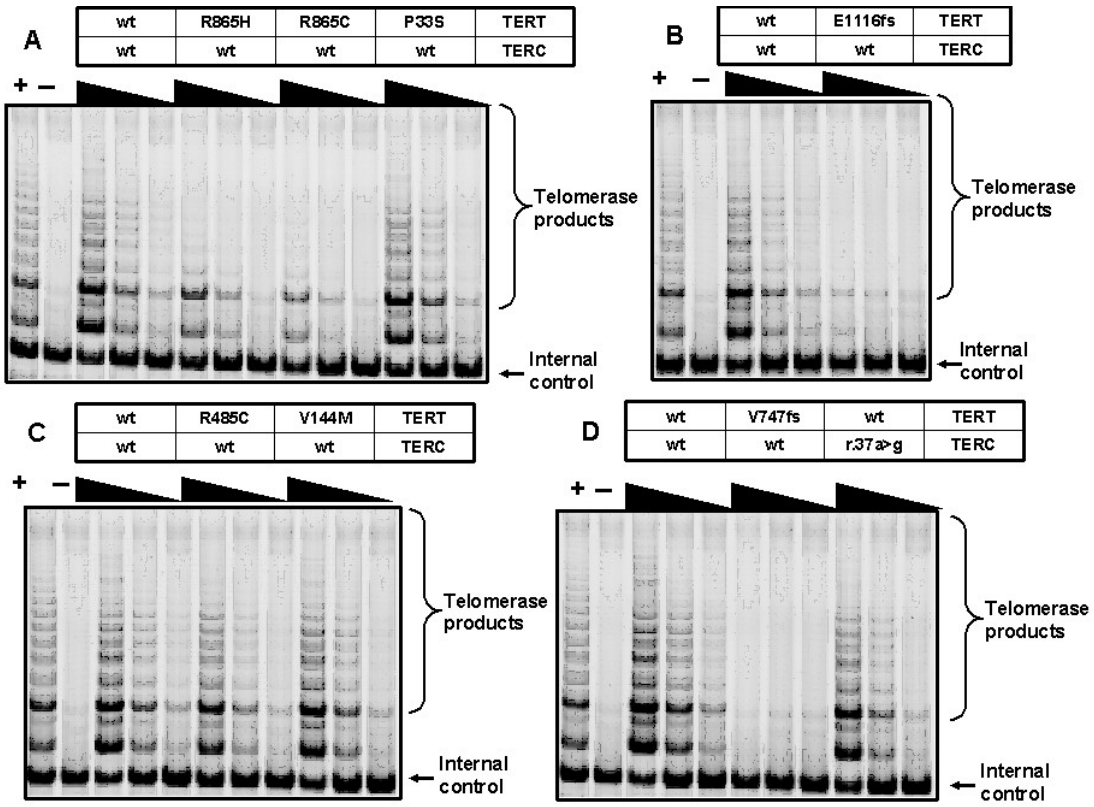

Figure 10
TRAP assays to measure telomerase activity for the seven tert mutations – five missense (A, C) and two deletions (B, D) - and the TERC mutation (D), as well. Plasmids encoding terc and tert were in vitro translated and transcribed and then tested for telomerase activity. Serial dilutions 1:5 were subjected to TRAP assay to ensure linear progression. Telomerase activity is calculated as the ratio of the intensity of the telomerase six bps ladder products to the internal control band intensity of each reaction. The relative activity for each sample is normalized to the wild type telomerase activity. In each experiment the positive control wild type TERC or TERT was cloned in the same vector as the mutant terc or tert. (+) sample is 250 cell equivalents of whole call lysates from a cell line positive for telomerase activity (H1299). (-) sample is a heat inactivated aliquot of the highest concentration of wild type telomerase transcription translation reaction prior to measuring the telomerase activity.

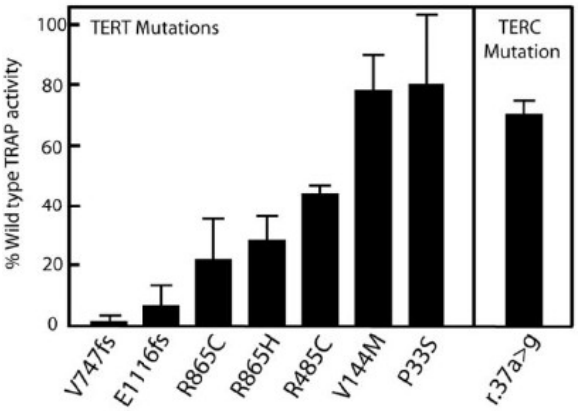

Figure 11
The relative telomerase activity for the seven tert mutations and the terc mutation compared to the wild type are shown. The relative activity of the mutant proteins and the mutant RNA varies from almost 0% to 100% of the wild type activity.

Mutations associated with an autosomal dominant pattern of inheritance can either be due to haploinsufficiency or to dominant negative mutations. To test whether our most severe mutation was either a dominant negative or a loss of function mutation, we designed a mixing experiment to distinguish the two possibilities. We co expressed wt tert and V747fsX tert in various ratios with wt terc and then process to TRAP assay. With the addition of the mutant tert to the wild type enzyme there is loss of net activity proportional to the amount of the mutant tert added in the reaction. This observation shows that the underlying mechanism for the disease caused by the V747fsX mutation is not a dominant negative type of action of the protein but haploinsufficiency (Figure 12).

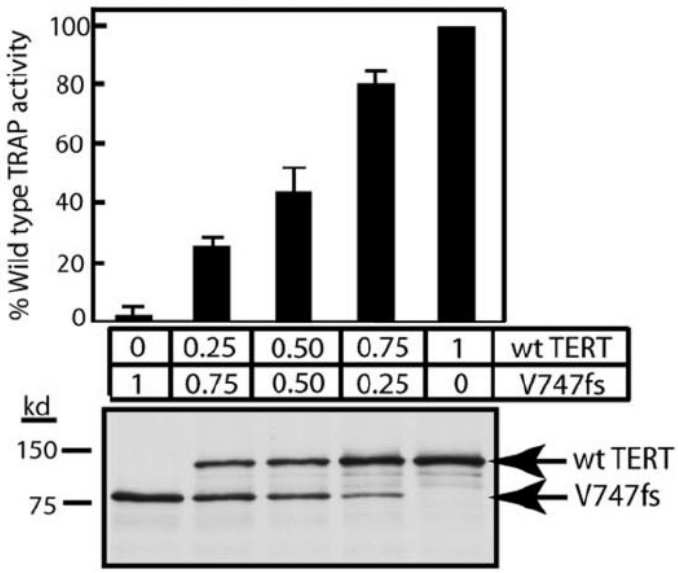

Figure 12
Wild type and V747fs tert mixing experiment to test the effect of the mutant protein on the wt one. The TnT reactions are run in parallel with reactions using [^{35}S] methionine to confirm the ratios of expression are the same as the ratios of plasmid added in the reactions. The V747fs tert is migrating faster in the acrylamide gel as the size is predicted to be smaller due to the premature termination codon.

In order to determine the effect of these mutations on the telomeres of our patients we wanted to measure their telomere length. An indirect way to measure the telomere length of the probands and their family members is the terminal restriction fragment (TRF) assay. The TRF assay is based on the lack of recognition sequences for

restriction enzymes in the telomeric repeats and the subtelomeric DNA. This way the length of the DNA fragments remaining after the genomic DNA has been digested with frequent cutters is representative to the sum of the telomeric repeats and the subtelomeric DNA size. To measure the size of these fragments the DNA is probed with radiolabeled oligos containing the TTAGGG repeat (Herbert et al. 2003) (Figure 13).

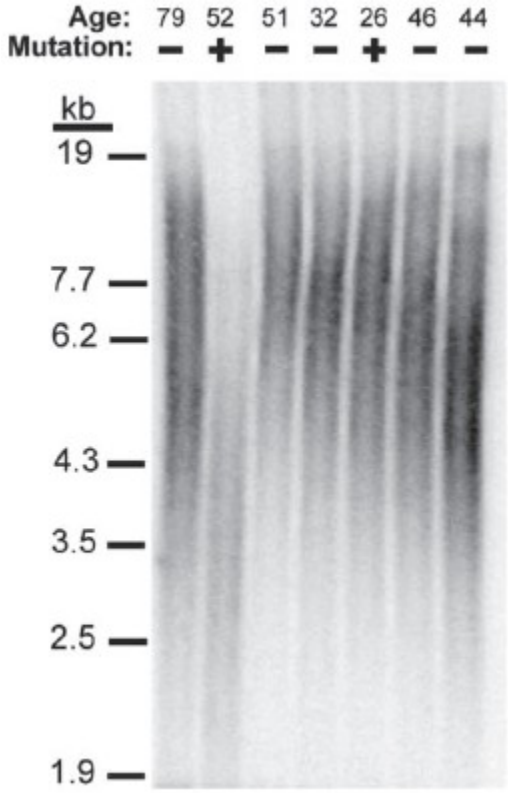

Figure 13
TRF analysis of genomic DNA from the Family 8. The DNA fragments produced by the digestion are resolved by gel eloctrophoresis, Southern blotted and plotted with a probe complimentary to the telomeric repeats. The smear showed in each lane is telomere specific. The individual 52year old is the proband of Family 8 that was diagnosed with IPF. He was found to have much shorter average telomere length than his age matched relative (51year old) or even elder member of his family (79year old) that were negative for the E1116fsX mutation. (+) and (-) indicate positive and negative for the family mutation respectively.

Telomere length was measured for all members of families with available genomic DNA. We compared the mean telomere length of carriers with age matched non-carriers within each family and found that it was shorter for the heterozygous ones. That was consistent for all mutations including the ones that did not affect the telomerase activity when tested with the in vitro TRAP assay. Combining the data from all families the carriers had shorter TRFs than non-carriers without a difference between affected and not affected carriers (Figure 14).

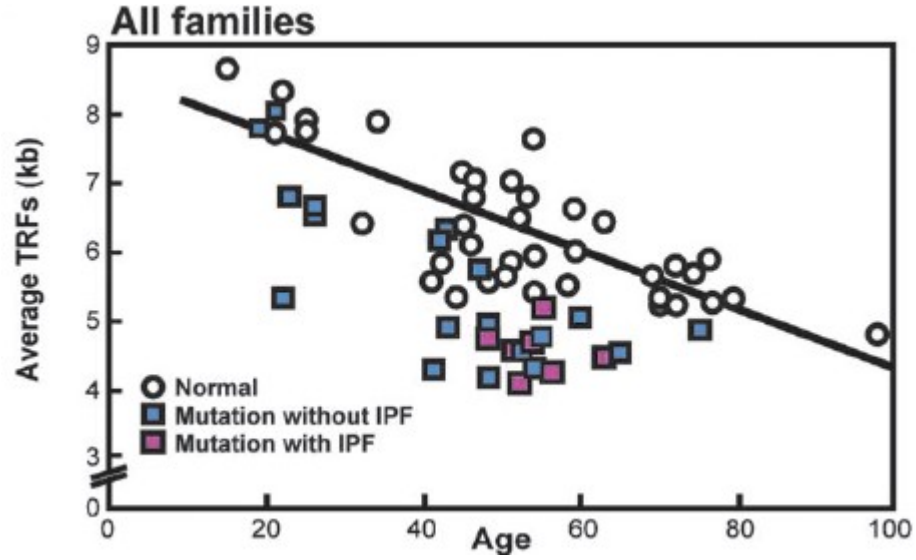

Figure14
The average TRFs of each individual is plotted against age. This figure is combining information from all families of the present study. The mean telomere length of asymptomatic carriers was similar to carriers with IPF and shorter than normal. The average telomere length is reducing with age as previously shown (Hastie et al. 1990) following a linear progression.

When we analysed the TRFs within each family we observed that the carriers with IPF had shorter telomeres than the asymptomatic ones. The same was noticed when we estimated the proportion of short TRFs and compared to age matched family members (Figure 15).

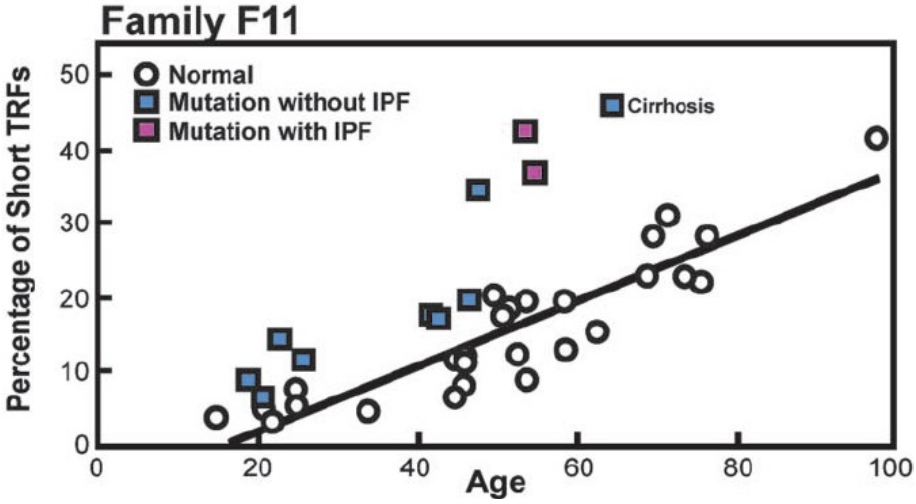

Figure 15
Percentage of Short TRFs plotted against the age of the members of Family11. Since it is the shortest telomeres and not the average telomeres that lead to telomere dysfunction (Hemann et al. 2006) we estimated the percentage of short telomeres for the members of Family 11. We arbitrarily define as short telomeres those between 1.9 and 4.3kbs. The percentage of short TRFs was calculated as the ratio of the relative signal intensity of the short telomeres area (1.9 to 4.3kbs) to the whole lane. The carriers have higher proportion of short telomeres compare to similar age healthy non carriers. Two individuals with R865H tert mutation and pulmonary fibrosis have the highest proportion of short telomeres. Another carrier has very short telomeres but not IPF. He had cirrhosis of unknown aetiology.

DISCUSSION

Using a non-biased whole genome linkage analysis we identified mutations in telomerase causing adult onset idiopathic pulmonary fibrosis. We screened 46 family probands and 44 individuals with sporadic IIPs. Seven mutations were identified in *TERT*, the gene encoding for the protein part of telomerase and one mutation in *TERC*, the gene encoding for the RNA part of the telomerase complex. The mutations in *TERT* are five missence mutations that involve highly conserved residues and two deletions that are predicted to result in premature stop codons. The mutation in TERC was a missence mutation previously described in a dyskeratosis congenita patient with severe aplastic anemia (Ly et al. 2005). In vitro reconstitution experiments measuring the telomerase activity of the mutated telomerase resulted in barely detectable activity for the two deletions and a range of 30 -100% compared to the wild type telomerase for the five missence *TERT* mutations and the transition *TERC* mutation. In addition we designed an experiment mixing different ratios of wild type and one *TERT* mutation predicted to be missing half the reverse transcriptase domain and subsequently measure the telomerase activity. The activity detected was proportional to the amount of the wild type protein added in the mixture suggesting a mechanism of haploinsufficiency. The telomere length of the affected individuals carrying a telomerase mutation was found to be significantly shorter than their asymptomatic carrier family members. These individuals had a higher proportion of short telomeres compared to their age matched family members. This way we provide evidence that idiopathic pulmonary fibrosis is a disease of telomere dysfunction.

In all cases the pulmonary fibrosis was developed in adult life distinguishing from previously reported familial forms of IIPs due to surfactant protein C (*SFPTC*) mutations (Thomas et al. 2002, Nogee et al. 2001). None of the individuals with *TERC* or *TERT* mutations have the characteristic triad or other clinical manifestations of DKC. None has been diagnosed for aplastic anemia although a mild to moderate anemia (hematocrit 23- 32%) was occasionally found among telomerase mutations carriers. Mild anemia could be considered as an additional characteristic of the disease. Furthermore osteoporosis/osteopenia was present in more than one third of affected with telomerase mutation and in half cases the diagnosis could not be attributed to steroid treatment, suggesting that it could be considered as an additional clinical indication of pulmonary fibrosis due to telomere shortening. A recent study though showed that osteoporosis was found not to be correlated to leukocyte telomere length in 2750 participants (Sanders et al. 2009).

All probands in this study fulfilled the criteria for IPF while there were telomerase mutatin carriers with unclassified pulmonary disease. In Family 71 the step sister of the proband was a smoker that presented with cough and dyspnea. The HRCT showed peripheral interstitial fibrosis of the lungs while the open lung biopsy specimen was read as UIP versus UIP with features of noncaseating granuloma. Since IPF is by far the most common clinicopathologic entity of the IIPs (ATS/ERS, 2002), it could be that telomerase mutation carriers are susceptible to a broad range of IIPs and not only to IPF.

The pattern of inheritance in all families was autosomal dominant with incomplete penetrance. The mutations were segregating with the phenotype. The incomplete penetrance of the disease may be due to the influence of environmental factors on a genetic predisposed individual.

Smoking is an independent risk factor for the development of IPF associated with the familial interstitial pneumonia (Steele at al. 2005). Cigarette smoking enhances

telomere shortening and this occurs in a dose response manner (Morla et al. 2006). In the present study the average age of death of the smokers was 10 years earlier than the non smokers.

The missence mutations that we identified in patients with IPF are mapped all across the entire tert protein. Two mutations, R865C and R865H, located in the C- motif of the Reverse Transcriptase domain indicate phenotype genotype relationship. None of the *TERT* mutations that we found has ever been reported in families with DKC or aplastic anemia. The r.37 a>g *TERC* mutation has already been described in a patient with DKC and aplastic anemia (Ly et al. 2005). The different phenotype resulting from the same mutation (pulmonary fibrosis or bone marrow failure) might be explained by environmental factors that trigger tissue specific responses (e.g. smoking for the lung). The proband of Family 61 that was heterozygous for the r.37 a>g *TERC* mutation was a smoker. His son, who was a carrier of the *TERC* mutation and smoker as well, presented with dyspnea and minimal scaring in HRCT.

In vitro activity of the mutant telomerase ranged from almost 0 to 100% compared to the wild type telomerase using the TRAP assay. The deletion mutations were close to 0% while the missence ones varied from 30 -100%. All of the missence mutations involved highly conserved residues including those that exhibited activity similar to the wild type in the TRAP assay. We showed that the carriers of the mutations had shorter telomeres than non carrier family members independently of the telomerase activity the mutant telomerase showed in vitro. Although in vivo the mutant telomerase could not maintain the telomere length and resulted in telomere shortening, the telomerase activity measured in vitro was not affected. This incongruity between in vivo and in vitro results reflects the limitations of the in vitro experiments. Since telomere length is attributed not only to telomerase activity but to other factors as replicative history and response to oxidative stress, as well (Garcia et al. 2007), we could speculate that some of these mutations might lead to telomere dysfunction compromising other functions of telomerase and not telomerase activity per se (Cong and Shay, 2008).

The V747fsX tert was shown to cause telomere dysfunction through haploinsufficiency. We mixed wild type tert and mutant tert mimicking the heterozygous state and showed that there was not enough active telomerase. The implications of this observation are very significant concerning putative therapeutic approaches.

There are currently various methods available for the measurement of the telomere length. TRF was the first technology developed to estimate the average telomere length in a cell population. It is a straight forward technique that gives a resolution of about 1kb. In Q- FISH and flow FISH telomeres are hybridized with complimentary PNA probes. The flow FISH has the advantage of hybridization in suspension and then perform flow cytometry. The resolution is two times higher than TRFs and we can distinguish cell subpopulations when staining with different antibodies. Single telomere length analysis (STELA) technique requires gel electrophoresis of the PCR amplified products and then Southern blot of the sample. Using chromosome specific primers for subtelomeric regions we can measure the length of single telomeres. Another method is Q-PCR where telomere repeats are specifically amplified and directly quantified. The Q-PCR estimates the average telomere length with a resolution as high as 76bps (Aubert and Lansdorp, 2008).

In our study we used TRFs to measure the telomere length of the participants and showed that PF is the result of telomere dysfunction. When we compared the telomere length of affected carriers with their relatives that were carrying the mutation but did

not have the disease we showed that the former ones tended to have shorter telomeres. In fact most of the mutation carriers who are younger (i.e. 45 years or younger) have telomeres length and percent short telomeres close to their normal control as seen within Family 11 (Figure 15). More specifically in Family 11 only four heterozygous individuals are age 50 or greater. Two have IPF and one has liver cirrhosis of unknown aetiology. The latter one has very short telomeres. Liver fibrosis of unknown cause has been reported in patients with DKC and IPF (Armanios et al. 2005).

A separate study using candidate gene approach showed that mutations in telomerase components lead to familial form of idiopathic pulmonary fibrosis in about 8% of cases. They identified five mutations in *TERT* and one mutation in *TERC*. None of the carriers had DKC while the family with the 98g>a *TERC* mutation had aplastic anemia. Using the flow FISH on peripheral- blood lymphocytes they found that carriers of the telomerase mutation either affected or unaffected had shorter telomeres than their age matched relatives. In addition, it was noted that the unaffected carriers were on average eleven years younger than the probands at the time of the diagnosis reflecting the age dependent onset of the disease. They proposed that pulmonary fibrosis is caused by telomere shortening on alveolar epithelium progenitor cells. Loss of these cells leads to inability of the epithelium to be replaced and the remodelling response appears as pulmonary fibrosis (Armanios et al. 2007). This independent study using different strategies than ours replicates part of our results and strengthens their significance.

Aside from the mutations reported in the present study and five more *TERT* mutations and a *TERC* mutation by Armanios et al. more telomerase mutations in patients with IPF have been described in subsequent studies. The Armanios group reported another *TERC* mutation (Alder et al. 2008) while six additional *TERT* mutations, five in familial and one in a sporadic case have been identified (Cronkhite et al. 2008). The *TERT* mutations in relation to the conserved domains of the telomerase protein are shown (Figure 16). None of these mutations has been described in patients with DKC. A V694M tert mutation identified in a family (Cronkhite et al. 2008) has been previously reported in an individual with aplastic anemia (Yamaguchi et al. 2005). On the other hand the *TERC* mutations are better associated with DKC and bone marrow failure than with IPF. Out of the four mutations described (37 a>g in the present study, 98 g>a by Armanios et al, 53-87 del by Marrone et al and 325g>t by Alder et al.) 37 a>g has already been described in a patient with DKC and aplastic anemia, 98g>a in a family with IPF and aplastic anemia (Garcia et al. 2007) and 53-87 del was identified in a family with pulmonary fibrosis and aplastic anemia (Marrone et al. 2007).

In an attempt to describe additional features of the pulmonary fibrosis telomere dysfunction syndrome, a group of individuals with IIPs and unexplained liver cirrhosis was identified concerning 3% of individuals with IIPs. In all cases these individuals have very short telomeres (Alder et al. 2008). It could be that the heterozygous for the R865H individual of Family 11 that had liver cirrhosis of unknown aetiology and very short telomeres falls within this cluster and did not develop the full phenotype by the time of his death by accident.

Osteoporosis/osteopenia was detected in half the IIPs affected individuals with telomerase mutations without though a significant difference in incidence compared to familial or sporadic IIPs without telomerase mutations (Cronkhite et al. 2008).

In a subsequent study granulomatous lung disease was detected in the proband of family who was found to be positive for a transition *TERT* mutation L1019F. In the

same study IPF was the most common diagnosis with 65% of telomerase mutation carriers meeting the criteria for IPF while other carriers were diagnosed with some type of unclassified pulmonary disease (Cronkhite et al. 2008). Combining with our observations carriers of these telomerase mutations are more susceptible to IIPs in general and not to one specific subtype (Cronkhite et al. 2008).

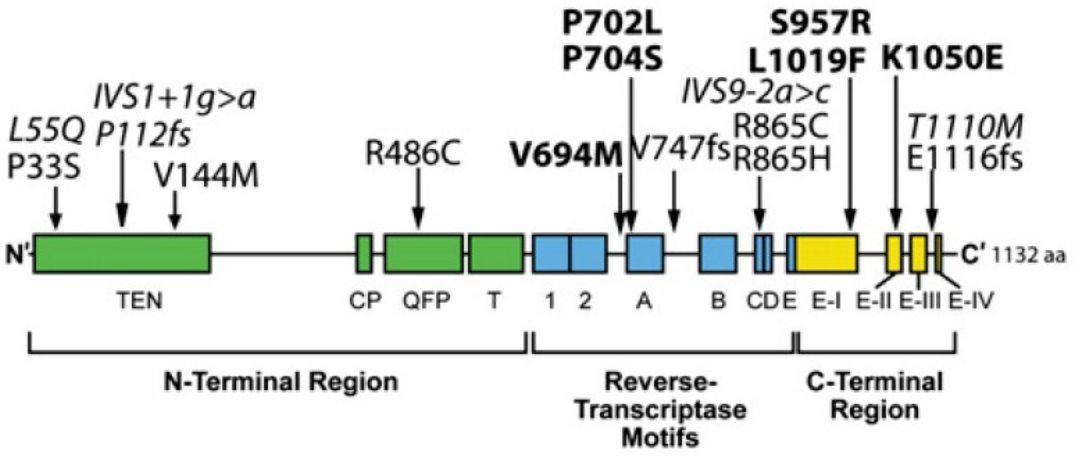

Figure 16
TERT mutations described in IPF sporadic and familial cases are shown in a schematic representation of the conserved areas of the tert protein. Mutations in roman are described in the currently presented study, in italics are described by Armanios et al and in boldface type by Cronkhite et al. Figure adapted by Cronkhite et al. 2008. Three independent observations in motif C of the RT domain suggest a genotype phenotype relationship. IVS9-2a>c mutation is located at a consensus splice junction leading to the skipping of exon 10 without distortion of the reading frame. The predicted mutant tert protein is missing the essential motif C (Armanios et al. 2007).

With the results presented in this study the group of diseases caused by telomere shortening expands to include pulmonary fibrosis suggesting telomerase pathway involved in the disease development. A plausible mechanism for the pathogenesis of pulmonary fibrosis may involve telomere shortening of a population of cells in the lung that have the ability to regenerate the tissue. This type of cells is recruited to respond to sequential injury caused by environmental factors like smoking. The enhanced proliferation of these cells in combination with the telomerase deficiency due to mutations results in shorter telomeres. The clinical phenotype is not present until the telomeres reach a critical short size driving the cells to replicative senescence and subsequently tissue dysfunction.
The identification of the type of cells that are responsible for the regeneration of the lung and the prevention of premature aging is of great interest. We could speculate that a subpopulation of type II alveolar cells is affected by the telomerase mutations since these cells have been shown to express telomerase (Driscoll et al. 2000). Since type II alveolar cells can transform into type I cells and repopulate the alveolar epithelium (Evans et al. 1975) their loss would result in inability of the lung to regenerate after injury. This is in agreement with the current hypothesis on IPF development due to *SFPTA2* mutations (Wang et al. 2009).

The adult onset of the disease might be due to the multiple rounds of cell division that the cells need to undergo before telomeres become critically short. The incomplete penetrance of the disease reflects the impact of the environmental factors or secondary 'hits' like susceptibility to fibrotic response in injury. The phenotypic diversity exhibited by telomerase mutations – e.g. pulmonary fibrosis vs. aplastic anemia- could be attributed to environmental factors as well as the nature of the mutation. Cigarette smoking causes oxidative damage that will result in shorter telomeres in the cells of the alveolar epithelium. That might explain why the predominant phenotype is lung disease in these cases. Thus there is a balance in telomere maintenance that when disturbed results in pathologic phenotype (Figure 17).
Understanding the pathogenesis of this fatal disease may lead to novel therapeutic approaches. In cases that the mode of action of the telomerase mutation is haploinsufficiency, as we have shown for the V747fsX tert mutation, enhancing telomerase expression in this subpopulation of cells could present new therapeutic approaches to this disorder. In other cases preventing telomere shortening might attenuate the progress of the disease.
Identifying telomerase mutations as the cause of adult onset pulmonary fibrosis could have an impact on the prognosis of affected individuals. In our study we could not predict the course of the disease or the survival according to mutations detected. Individuals with IPF do not respond to current immunosuppressive therapies (Meltzer and Noble. 2008). The telomerase mutation is an additional feature to exclude other cases of pulmonary fibrosis in cases that the diagnosis is not so clear. These patients might benefit more by new treatments that are still in experimental level. Finally genotyping unaffected family members of affected individuals is controversial as there are no current regiments to prevent the disease and the penetrance of the disease is variable.

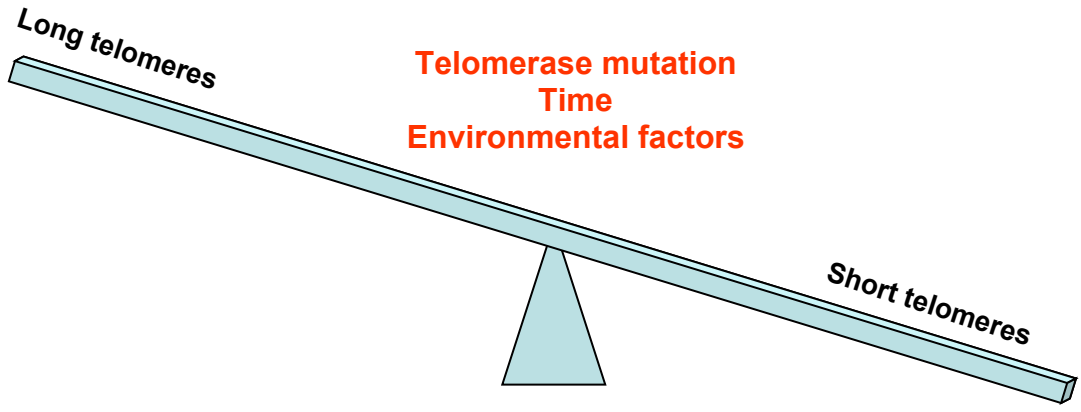

Figure 17
A proposed model describing the pathogenesis of pulmonary fibrosis. The presence of telomerase mutations and the number of cell divisions in a cellular population tip the balance towards shorter telomeres. Environmental factors that cause oxidative stress like cigarette smoking exacerbate the attrition of telomeres in cells in the lung. This way we have impaired tissue response to injury and lung disease phenotype.

Acknowledgments:
I would like to thank the affected individuals and their families for participation in the study; Dr. Christine Garcia; all the members of her laboratory as well as Joseph Garcia for helpful advice; Jerry Shay laboratory members for their assistance on the TRAP assays; the Director of the graduate programme Dr. Boumpas, and the graduate program for introducing me to science; the secretary of the graduate program Mary Adamaki; Dr. Vassilis Zannis for guidance; the advisory committee; the seven member committee; my parents Chrisoula and Dimitris Tsakiris and my spiritual father Fr. Asterios Kaimakis for encouragement and support; and finally my husband Justin M. Hicks for his understanding and love.

REFERENCES

Abecasis, G. R., Cherny, S. S., Cookson, W. O., and Cardon, L. R. (2002). Merlin rapid analysis of dense genetic maps using sparse gene flow trees. Nat Genet. *30*, 97–101.

Alder, J. K., Chen, J. J., Lancaster, L., Danoff, S., Su, S. C., Cogan, J. D., Vulto, I., Xie, M., Qi, X., Tuder, R. M., Phillips, J. A., Lansdorp, P. M., Loyd, J. E., and Armanios, M. Y. (2008). Short telomeres are a risk factor for idiopathic pulmonary fibrosis. Proc Natl Acad Sci U S A. *105*, 13051-6.

American Thoracic Society/European Respiratory Society. (2002). American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. Am J Respir Crit Care Med. *165*, 277–304.

American Thoracic Society (ATS), and the European Respiratory Society (ERS). (2000). Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). Am J Respir Crit Care Med. *161*, 646–664.

Armanios, M., Chen, J. L., Chang, Y. P., Brodsky, R. A., Hawkins, A., Griffin, C. A., Eshleman, J. R., Cohen, A. R., Chakravarti, A., Hamosh, A., Greider, C. W. (2005). Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. Proc Natl Acad Sci USA. *102*, 15960–15964.

Armanios, M., Chen, J. L., Cogan, J. D., Alder, K. J., Ingersoll, R. G., Markin, C., Lawson, W. E., Xie, M., Vulto, I., Phillips, J. A., Lansdorp, P. M., Greider, C. W., and Loyd, J. E. (2007). Telomerase Mutations in Families with Idiopathic Pulmonary Fibrosis. N Engl J Med. *356*,1317-26.

Aubert, G., and Lansdorp P. M. (2008). Telomeres and aging. Physiol Rev. *88*, 557-79.

Autexier, C., and Neal, F. L. (2006). The Structure and Function of Telomerase Reverse Transcriptase. Annu Rev Biochem. *75*, 493–517.

Azuma, A., Nukiwa, T., Tsuboi, E., Suga, M., Abe, S., Nakata, K., Taguchi, Y., Nagai, S., Itoh, H., Ohi, M., Sato, A., Kudoh, S., and Raghu, G. (2005). Double-blind, Placebo-controlled Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. Am J Respir Crit Care Med. *171*, 1040–1047.

Banik, S. R., Guo, C., Smith, A. C., Margolis, S. S., Richardson, D. A., Tirado, C. A., and Counter, C. M. (2002). C-Terminal Regions of the Human Telomerase Catalytic Subunit Essential for In Vivo Enzyme Activity. Mol Cell Biol. *22*, 6234–6246.

Bjoraker, J. A., Ryu, J. H., Edwin M. K., Myers J. L., Tazelaar H. D., Schroeder D. R., and Offord K. P.(1998). Prognostic Significance of Histopathologic Subsets in Idiopathic Pulmonary Fibrosis. Am J Respir Crit Care Med *157*, 199–203.

Blasco, M. A. (2005). Mice with bad ends: mouse models for the study of telomeres and telomerase in cancer and aging. EMBO J. *24*, 1095–1103.

Blasco, M. A. (2007). Telomere length, stem cells and aging. Nat Chem Biol.*10*, 640-9.

Calado, R. T., Regal, J. A., Hills, M., Yewdell, W. T., Dalmazzo, L. F., Zago, M. A., Lansdorp, P. M., Hogge, D., Chanock, S. J., Estey, E. H., Falcão, R. P., and Young, N. S. (2009). Constitutional hypomorphic telomerase mutations in patients with acute myeloid leukemia. Proc Natl Acad Sci U S A. *106*, 1187-92.

Chen, J. L., Blasco, M. A., and Greider, C. W. (2000). Secondary structure of vertebrate telomerase RNA. Cell. *100*, 503-514.

Chen, J. L., and Greider, C. W. (2003). Template boundary definition in mammalian telomerase. Genes Dev. *17*, 2747–2752.

Cohen, S. B., Graham, M. E., Lovrecz, G. O., Bache, N., Robinson, P. J., and Reddel, R. R.(2007). Protein composition of catalytically active human telomerase from immortal cells. Science. *315*, 1850-3

Cong, Y., and Shay, J. W. (2008). Actions of human telomerase beyond telomeres. Cell Res. *7*, 725-32.

Cronkhite, J. T., Xing, C., Raghu, G., Chin, K. M., Torres, F., Rosenblatt, R. L., and Garcia, C. K. (2008). Telomere shortening in familial and sporadic pulmonary fibrosis. Am J Respir Crit Care Med. *178*, 729-37.

Demedts, M., Behr, J., Buhl, R., Costabel, U., Dekhuijzen, R., Jansen, H. M., MacNee, W., Thomeer, M., Wallaert, B., and Laurent, F. (2005). High-dose acetylcysteine in idiopathic pulmonary fibrosis. N Engl J Med. *353*, 2229–2242.

Dokal, I. (2000). Dyskeratosis congenita in all its forms. Br J Haematol. *110*, 768–779.

Driscoll, B., Buckley, S., Bui, K. C., Anderson, K. D., and Warburton, D. (2000). Telomerase in alveolar epithelial development and repair. Am. J. Physiol. Lung Cell. Mol. Physiol. *279*, 1191–1198.

Evans M. J., Cabral L. J., Stephens R. J. and Freeman G. (1975). Transformation of alveolar Type 2 cells to Type 1 cells following exposure to NO2. Exp Mol Pathol. *22*, 142-50.

Garcia, C. K., Wright, W. E., and Shay, J. W. (2007). Human diseases of telomerase dysfunction: insights into tissue aging. Nucleic Acids Res. *35*, 7406-16.

Gillis, A. J., Schuller, A. P., and Skordalakes, E. (2008). Structure of the Tribolium castaneum telomerase catalytic subunit TERT. Nature. *455*, 633-7.

Griffiths, A. J. F. (2002). An intoduction to genetic analysis. (Freeman, NY). 7th Edition, 258-259.

Gross, T. J., and Hunninghake, G. W. (2001). Idiopathic Pulmonary Fibrosis. N Engl J Med *345*, 517–525.

Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D. K., and Allshire, R. C. (1990). Telomere reduction in human colorectal carcinoma and with ageing. Nature. *346*, 866-8.

Heiss, N. S., Knight, S. W., Vulliamy, T. J., Klauck, S. M., Wiemann, S., Mason, P. J., Poustka, A., and Dokal, I. (1998). X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nature Genet. *19*, 32–38.

Hemann, M. T., Strong, M. A., Hao, L. Y., and Greider, C. W. (2001). The shortest telomere, not average telomere length, is critical for cell viability and chromosome stability. Cell. *107*, 67–77.

Herbert, B-S., Shay, J. W., and Wright, W. E. (2003). Current Protocols in Cell Biology (Wiley, Hoboken, NJ), Vol 18.6.1–18.6.20.

Herbert, B., Hochreiter, A. E., Wright, W. E., and Shay, J. S. (2006). Nonradioactive detection of telomerase activity using the telomeric repeat amplification protocol. Nat Protocols. *1*, 1583–90.

Jacobs, S. A., Podell, E. R., and Cech, T. R. (2006). Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. Nat Struct Mol Biol. *13*, 218-25.

Katzenstein, A. L., and Myers, J. L. (1998). Idiopathic pulmonary fibrosis: clinical relevance of pathologic classification. Am J Respir Crit Care Med. *157*, 1301–1315.

Kim, N. W., Piatyszek, M. A., Prowse, K. R., Harley, C. B., West, M. D., Ho, P. L., Coviello, G. M., Wright, W. E., Weinrich, S. L., and Shay, J. W. (1994). Specific association of human telomerase activity with immortal cells and cancer. Science. *266*, 2011-5.

Kolquist, K. A., Ellisen, L. W., Counter, C. M., Meyerson, M., Tan, L. K., Weinberg, R. A., Haber, D. A., and Gerald, W. L. (1998). Expression of TERT in early premalignant lesions and a subset of cells in normal tissues. Nat Genet. *19*, 182-6.

Kong, A., and Cox, N. J. (1997). Allele-sharing models: LOD scores and accurate linkage tests. Am J Hum Genet. *61*, 1179–1188.

Ly, H., Schertzer, M., Jastaniah, W., Davis, J., Yong, S. L., Ouyang, Q., Blackburn, E. H., Parslow, T. G., and Lansdorp, P. M. (2005). Identification and functional characterization of 2 variant alleles of the telomerase RNA template gene (TERC) in a patient with dyskeratosis congenita. Blood. *106*, 1246–1252.

Marrone, A., Sokhal, P., Walne, A., Beswick, R., Kirwan, M., Killick, S., Williams, M., Marsh, J., Vulliamy, T., and Dokal, I. (2007). Functional characterization of novel telomerase RNA (TERC) mutations in patients with diverse clinical and pathological presentations. Haematologica. *92*, 1013-20.

Marshall, R. P., Puddicombe, A., Cookson, W. O., and Laurent, G. J. (2000). Adult familial cryptogenic fibrosing alveolitis in the United Kingdom.Thorax. *55*, 143–146.

Meltzer, E. B., and Noble, P. W. (2008).Idiopathic pulmonary fibrosis. Orphanet Journal of Rare Diseases. *26*, 3:8.

Morla, M., Busquets, X., Pons, J., Sauleda, J, MacNee, W., and Agusti, A. G. (2006). Telomere shortening in smokers with and without COPD. Eur Respir J. *27*,525–528.

Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., Harley, C. B., and Cech, T. R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. Science *277*, 955–959.

Nogee, L. M., Dunbar, A. E., Wert, S. E., Askin, F., Hamvas, A., and Whitsett, J. A. (2001). A mutation in the surfactant protein C gene associated with familial interstitial lung disease. *344*, 573-9.

Raghu, G., Weycker, D., Edelsberg, J., Bradford, Z. W., and Oster, G. (2006). Incidence and Prevalence of Idiopathic Pulmonary Fibrosis. Am J Respir Crit Care Med. *174*, 810–816.

Schageman, J. J., Horton, C. J., Niu, S., Garner, H. R., and Pertsemlidis, A. (2004). ELXR: a resource for rapid exon-directed sequence analysis. Genome Biol. *5*, R36.

Steele, M. P., Speer, M. C., Loyd, J. E., Brown, K. K., Herron, A., Slifer, S. H., Burch, L. H., Wahidi, M. M., Phillips, J. A., and Sporn, T. A. (2005). Clinical and pathologic features of familial interstitial pneumonia. Am J Respir Crit Care Med. *172*, 1146–1152.

Thomas, A. Q., Lane, K., Phillips, J., III, Prince, M., Markin, C., Speer, M., Schwartz, D. A., Gaddipati, R., Marney, A., and Johnson, J. (2002). 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. *165*, 1322–1328.

Tsakiri, K. D., Cronkhite, J. T., Kuan, P. J., Xing, C., Raghu, G., Weissler, J. C., Rosenblatt, R. L., Shay, J. W., and Garcia C. K. (2007). Adult-onset pulmonary fibrosis caused by mutations in telomerase. Proc Natl Acad Sci U S A. *104*, 7552-7.

Vulliamy, T., Marrone, A., Goldman, F., Dearlove, A., Bessler, M., Mason, P. J., and Dokal, I. (2001). The RNA component of telomerase is mutated in autosomal dominant dyskeratosis congenita. Nature. *413*, 432–435.

Vulliamy, T., Marrone, A., Dokal, I., and Mason, P. J. (2002). Association between aplastic anaemia and mutations in telomerase RNA. Lancet. *359*, 2168-70.

Vulliamy, T., Beswick, R., Kirwan, M., Marrone, A., Digweed, M., Walne, A. J., and Dokal, I. (2008). Mutations in the telomerase component NHP2 cause the premature ageing syndrome dyskeratosis congenital. Proc Natl Acad Sci USA. *105*, 8073–8078.

Walne, A. J., Vulliamy, T., Marrone, A., Beswick, R., Kirwan, M., Masunari, Y., Al-Qurashi, F., Aljurf, M., and Dokal, I. (2007). Genetic heterogeneity in autosomal recessive dyskeratosis congenita with one subtype due to mutations in the telomerase-associated protein NOP10. Hum Mol Genet. *16*, 1619-29.

Walne, A. J., Vulliamy, T., Beswick, R., Kirwan, M., and Dokal, I. (2008). TINF2 mutations result in very short telomeres: analysis of a large cohort of patients with dyskeratosis congenita and related bone marrow failure syndromes. Blood. *112*, 3594–3600.

Wang, Y., Kuan, P. J., Xing, C., Cronkhite, J. T., Torres, F., Rosenblatt, R.L., DiMaio, J. M., Kinch, L. N., Grishin, N. V., and Garcia, C. K. (2009). Genetic Defects in Surfactant Protein A2 Are Associated with Pulmonary Fibrosis and Lung Cancer. Am J Hum Genet. *84*, 52–59.

Weinrich, S. L., Pruzan, R., Ma, L., Ouellette, M., Tesmer, V. M., Holt, S. E., Bodnar, A. G., Lichtsteiner, S., Kim, N. W., Trager, J. B., Taylor, R. D., Carlos, R., Andrews, W. H., Wright, W. E., Shay, J. W., Harley, C. B., and Morin, G. B.(1997). Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. Nature Genetics. *17*, 498-502.

Yamaguchi, H., Calado, R. T., Ly, H., Kajigaya, S., Baerlocher, G. M., Chanock, S. J., Lansdorp, P. M., and Young, N. S. (2005). Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. N Engl J Med *352*, 1413–1424.