"Pathogenesis and Clinical Manifestations of Pulmonary Fibrosis in Scleroderma"

Introduction

Scleroderma (Systemic Sclerosis; SSc) is a multifaceted connective disorder with a complex array of disease manifestations and clinical outcomes. Lung disease represents one of the most common and more serious complications in SSc. There are two main areas where the lung could be damaged: the pulmonary vasculature (presenting as pulmonary arterial hypertension) and scarring and fibrosis of the lung tissue itself. The extent and severity of the lung fibrosis can be variable, and can range from a mild, interstitial involvement to a rapidly progressive scarring disorder that is associated to severe alterations of the lung architecture, causing declining lung function and leading to respiratory failure. As this deadly complication has been more intensely studied and better understood, albeit still incompletely, it has become evident that other characteristics of the disease tend to coexist and may be used as predictor factors for the less favourable and aggressive type of lung disease. These include shorter time interval between the onset of the skin disease, the severity of lung involvement, and the presence of cardiac disease. Although the final mechanisms leading to pulmonary fibrosis remain to be fully characterised, it is clear that in the process leading to scarring of the lung tissues, many factors interact. We believe that the patient's genetic background is one of the key factors that impact upon susceptibility, severity and progression and it is likely that patient genetics influence the conditions that define the outcome from the initial damage into chronic inflammation and tissue fibrosis.

Overall objective of programme

The objective of this research is to identify and characterise the role of patient genetics in SSc, and to determine how different genetic attributes impact upon the susceptibility to SSc, the rate of progression and organ involvement. In the long-term we aim to utilise our understanding of genetics to improve patient management and outcome. Our main aims are to:

- Identify the genetic polymorphisms associated with risk of developing pulmonary fibrosis in SSc patients
- Determine how these genetic variations influence disease susceptibility, progression and outcome
- Assess whether these genetic differences are useful clinical biomarkers for SSc
- Investigate genetic targets as candidates for novel approaches to therapy

This work will provide critical insights into the mechanisms of lung fibrosis, and hopefully permit early identification of cases most likely to progress before major lung fibrosis has developed. This will provide real potential for improving the quality of health care and effective treatments for SSc.

Progress so far

There have been four areas of progress over the last 12 months:

1. Genome Wide Association study (GWAS) in SSc

We have used the powerful approach of GWAS, to explore over one million genetic variations in the genes of patients with SSc. These patients were selected using very stringent inclusion criteria to obtain a pristine, clinically homogeneous populations. The analysis of the results showed clear associations with different disease subsets, lung fibrosis and autoantibody profiles. These studies have highlighted several polymorphisms with strong associations with SSc and in particular with the group of patients with a specific tissue type and the presence of anti-topoisomerase. We have also conducted a confirmatory analysis of the 100 top genetic associations found in the original GWAS in an extended

group which consisted of 500 new SSc patient and a control cohort. This essential study, successfully confirmed the association found in 24 of the genetic variations we investigated. The second set also confirmed data from the discovery cohort identifying the importance of the HLA region in the predisposition to the disease. We are now embarking upon a fine search of this region in order to define the precise region of DNA responsible for the association.

2. **Candidate gene analysis in SSc**

We have performed several candidate gene studies in order to assess the presence of genetic differences in genes that we believe are involved in pulmonary fibrosis.

We used HapMap data to tag the majority of known common SNPs in seven biologically significant candidate genes in SSc: *ALK1(*19
SNPs), *ENG* (17), *SMAD2* (47), *SMAD4* (15), *TGFB1* (6), *TGFBR1* (43), *TGFBR2*(42). We subsequently used TaqMan SNP assays to analyse 55 Tag SNPs in 465 UK Caucasian patients with SSc and 2270 UK female controls (Breast Cancer Association Consortium). The SSc group was selected using stringent clinical criteria and follow-up data in order to obtain a clearly defined SSc autoantibody and clinical profiles. We found a weak association with a rare TGFBR1 SNP rs13289899 and we identified that this association was related to the Scl70 positive subset of patients within this group.

Other candidate genes that are currently under investigation are KCN5, STAT4 and NKX2-5. These polymorphisms were assessed in 500 patients and 300 controls. Results so far have confirmed the association of polymorphisms in the STAT4 gene with the presence of SSc and in particular with the diffuse form of the disease. These results further support the important role of the genes involved in the immune response in the susceptibility to SSc. The comprehensive analysis of SNPs in NKX2-5 gene showed a significant association of 2 polymorphisms with the presence of lung fibrosis. Functional studies are now underway to fully explore and define the role of the NKX2-5 polymorphisms on the gene function/regulation in SSc. The analysis of one polymorphism in the KCNA5 gene (which produces the important potassium channel protein KV1.5) failed to demonstrate an association with the disease although in patients with a particular genotype, there is a tendency towards a protective effect on the development of pulmonary hypertension. This initial observation needs to be further and fully investigated within an extended SSc disease cohort.

3. **Gene expression profiling in SSc and IPF - Common or unique genetic variations**

In this aim we have continued gene expression profiling studies on cultured pulmonary fibroblasts derived from SSc patients in order to determine the functional significance of differentially expressed genes. A comparison between SSc lung fibroblasts, those derived form idiopathic pulmonary fibrosis and control normal cells have revealed common mechanism(s) that drive pulmonary fibrosis and have highlighted a number of primary disease targets that we are pursuing. We are studying the expression of the proteins coded by these genes, and using methods to inhibit and suppress gene activity in order to assess how they influence fibroblast function in scarring and fibrosis.

The Future

Although this work is still at a relatively early stage, it is already producing exciting results and the last 12 months has been a very productive time. By defining the key genes and proteins that are altered in scleroderma lung disease we will be able to better understand the factors that determine whether lung fibrosis progresses or remains stable. We will also better understand the likely pathways to blocking fibrosis and the links with other aspects of the disease. Our novel GWAS approach, using small (~100) highly clinically and biochemically defined group of patients has been shown to be a valid and powerful strategy as it has confirmed previous GWAS studies with larger cohorts. More importantly, we had

identified 20 new and novel genetic areas of association, most of which have not been previously implicated in SSc. These findings open up a whole new era in our study of the genetics of SSc and the identification of disease-related genes. Moreover, our “candidate gene approach”, based on our increasing understanding of the clinical condition and disease pathogenesis is equally valuable and complementary, providing an additional and more extensive assessment of the genetic areas/genes highlighted on the gene expression studies with a potential explanation for the observed expression differences. Interestingly several of the genes highlighted in the profiling studies do not appear on the GWAS maybe due to incomplete coverage of the genome provided by the current GWAS platforms, or more complex genetic associations not identified by GWAS technology

Therefore we believe that our three themed approach to study of the genetics of SSc is valid, robust and will eventually provide not only strong biologic markers for diagnosis and prognosis of the disease, but also useful and effective therapeutic candidates.

Christopher Denton
Athol Wells
David Abraham
Elizabeth Renzoni
Carmen Fonseca
Panagiotis Pantelidis

November 2010