Samenvatting proefschrift van Begoña Diosdado Calvo

"The puzzle of coeliac disease: pieces of the molecular pathogenesis"

Coeliac disease (CD) is a genetic disorder in which environmental factors and multiple genes contribute to its development. The main environmental factor that triggers CD is a protein called gluten. This protein is present in common cereals such as wheat, barley and rye, and allows us to process their derivate products such as bread or pasta. The only known genetic factors for CD are the genes located in chromosome 6 that encode for the HLA-DQ2/DQ8 molecules, and the MYO9B gene located on chromosome 19. Clinical symptoms include a wide rage of features that vary from typical diarrhea and malabsorption (or failure to thrive in children), to fatigue, anaemia, weight loss or no symptoms at all. Currently, the only therapy consists of a strict, life-long, gluten-free diet.

The disease starts when a susceptible individual ingests gluten and this enters his/her body through the mucosa of the small intestine. There, gluten is able to activate an abnormal immune response that is going to lead to consecutive, well-established changes of the intestinal mucosa. These are divided into three stages: lymphocytosis (Marsh I stage), crypt hyperplasia (Marsh II) and villous atrophy (Marsh III). Over the past few years, our knowledge about the pathogenesis of CD has increased substantially. We have long known that gluten could activate both adaptive and innate responses. However, further characterization of the immune response and the mechanisms behind tissue remodelling are still poorly understood. The research performed in this project provided new insights into the molecular mechanisms of both the adaptive and innate immune response, and the molecular mechanisms that drive villous atrophy. Consequently, uncovering new aspects of the patho-physiology of the disease has directed our search for candidate genes that predispose to CD.

We first aimed to identify the transcriptional changes that occur in the intestine of CD individuals who show villous atrophy (Marsh III) compared to healthy controls. We compared the expression profiles of biopsies from 15 well-characterized Marsh III individuals and seven controls in a microarray experiment that encompassed 19,500 genes. We reported a set of 109 genes involved in the long-term tissue destruction. Many of these genes had a function in proliferation and differentiation pathways and might be involved in the homeostasis of the crypt and villi. This was further supported by comparing the expression profiles from Marsh III patients refraining from gluten to Marsh III patients ingesting gluten. Our analysis yielded a second set of 120 genes that revealed a decreased mitotic activity in Marsh III patients after gluten withdrawal. Based on these results we proposed that the separate processes of crypt hyperplasia and villous atrophy are molecularly intertwined and we hypothesized that a deficiency of the cells in the crypts in completing their terminal differentiation program and migrating towards the villous was responsible for the formation of hyperplasic crypts and the subsequent villous atrophy.

A second microarray experiment was carried out to specifically determine the transcriptional activity of the genes related to the immune response during mucosal renewal. We selected a

total of 66 biopsies from CD individuals in various stages of remission and from controls, which enabled us to set up a virtual time course experiment. This allowed us to determine immune-related differentially expressed genes and the dynamics of their pattern of expression during tissue regeneration. The results revealed that gene expression profiles from the differentially expressed genes showed a gradual increase or decrease from the moment the treatment was started to the point of full recovery. It was interesting that this observation did not correlate with the discrete histological (Marsh) stages used in diagnosing the disease. Thus, despite the morphological differences, Marsh I and Marsh II individuals cannot be distinguished at a molecular level from Marsh III patients. Hence, we proposed that all patients should be considered "coeliacs" as long as they respond to a gluten-free diet.

The classification of the genes within the adaptive and innate pathways showed a comparable contribution of both pathways in response to gluten withdrawal. Our results suggested that interferon gamma is able to maintain the chronicity of the inflammatory response by inducing the transcription of chemokines and cytokines that attract and further activate the Th1 immune response through the JAK-STAT1 pathway. In relation to the innate immune response, our expression studies showed several indicators of neutrophils being recruited into the lamina propria of CD patients. This observation was further confirmed by immunohistochemistry using neutrophil-specific antibodies in both Marsh III and completely recovered (Marsh 0) CD individuals. This result suggests that CD patients have a permanent state of activation of the innate immune response in the lamina propria, regardless of whether they have histological damage or exposure to dietary gluten. We speculated that the enhanced neutrophil infiltration might reflect an underlying effect of the MYO9B gene. The MYO9B might be involved in the impairment of the intestinal barrier and, hence, lead to an increased permeability as seen in CD individuals.

The integration of the gene expression data with known linkage data, combined with knowledge on the pathogenesis of the disease, led us to two functional candidate genes that were subsequently tested in genetic association studies. The prolyl-endopeptidase (PREP) enzyme, located under the 6q21-22-linkage peak, is able to cleave proline-rich gluten peptides. We hypothesized that an altered PREP activity could be responsible for an inefficient breakdown of gluten peptides, which could result in the accumulation of immunostimulatory gluten peptides and thereby play a role in the breakdown of tolerance to gluten. To investigate the role of the PREP gene as a primary factor in CD, we conducted sequence analyses, genetic association studies, and determined the PREP enzyme activity in biopsies from CD patients and controls. Our results from the genetic association studies and the activity determinations indicated that PREP is not a causative gene in the Dutch CD population. Last, we carried out genetic association studies on the positional and functional candidate signal transducer and activator of transcription 1 (STAT-1) gene, located under the 2q linkage peak. This gene encodes for one of the main down-stream modulators of interferon gamma. We speculated that any genetic change that alters its transcriptional activity could explain the detrimental immune response seen in CD. Again, no evidence for positive association was observed in the Dutch CD population.

The gene expression studies performed in this project on the biopsies of CD patients and controls have led to the identification of novel pathways involved in the pathogenesis of CD. This new knowledge will help the search for susceptibility genes in CD. Moreover, the integration of the genetic and genomic information will be key to developing new diagnostic tools for identifying individuals at risk, and to finding new avenues for therapeutic intervention in coeliac disease.

21 March 2006