Ghent University - Faculty of Sciences
Flanders Interuniversity Institute for Biotechnology
Department for Molecular Biomedical Research

Transcriptome Profiling and Genetic Analysis to Identify Susceptibility Genes for Crohn's Disease

Debby LAUKENS

Gent, June 2006

Thesis submitted in partial fulfilment of the requirements for the degree of

DOCTOR IN SCIENCES: BIOTECHNOLOGY

Promoters: Prof. Erik Remaut and Dr. Pieter Rottiers

Examination commission

Chairwoman

Prof. Marcelle Holsters[1]

Reading committee

Prof. Erik Remaut[2]

Dr. Pieter Rottiers[2]

Prof. Martine De Vos[3]

Prof. Dirk Elewaut[4]

Prof. Jurgen Del-Favero[5]

Other members

Prof. Lothar Steidler[6]

Prof. Filip De Keyser[4]

Dr. Pieter Demetter[7]

[1]Department of Plant Systems Biology, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium

[2]Department for Molecular Biomedical Research, Ghent University, VIB, Gent, Belgium

[3]Department of Gastroenterology, Ghent University Hospital, Gent, Belgium

[4]Department of Rheumatology, Ghent University Hospital, Gent, Belgium

[5]Department of Molecular Genetics, University of Antwerp, VIB, Antwerpen, Belgium

[6]Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland

[7]Department of Pathology, Erasmus University Hospital, Brussel, Belgium

Scope of the project 1

Section I

Chapter 1 7

Introduction to the molecular genetics of Crohn's disease

Chapter 2 39

A functional promoter polymorphism in *CARD15*
Manuscript in preparation

Chapter 3 49

Reduced metallothionein expression in colonic Crohn's disease: evidence for *MTF1* as a new disease-modifying gene
Submitted to Gastroenterology

Chapter 4 73

Human metallothionein expression under normal and pathological conditions: mechanisms of gene-regulation
Review in preparation

Section II

Chapter 5 101

Spondyloarthropathy as a model for early Crohn's disease

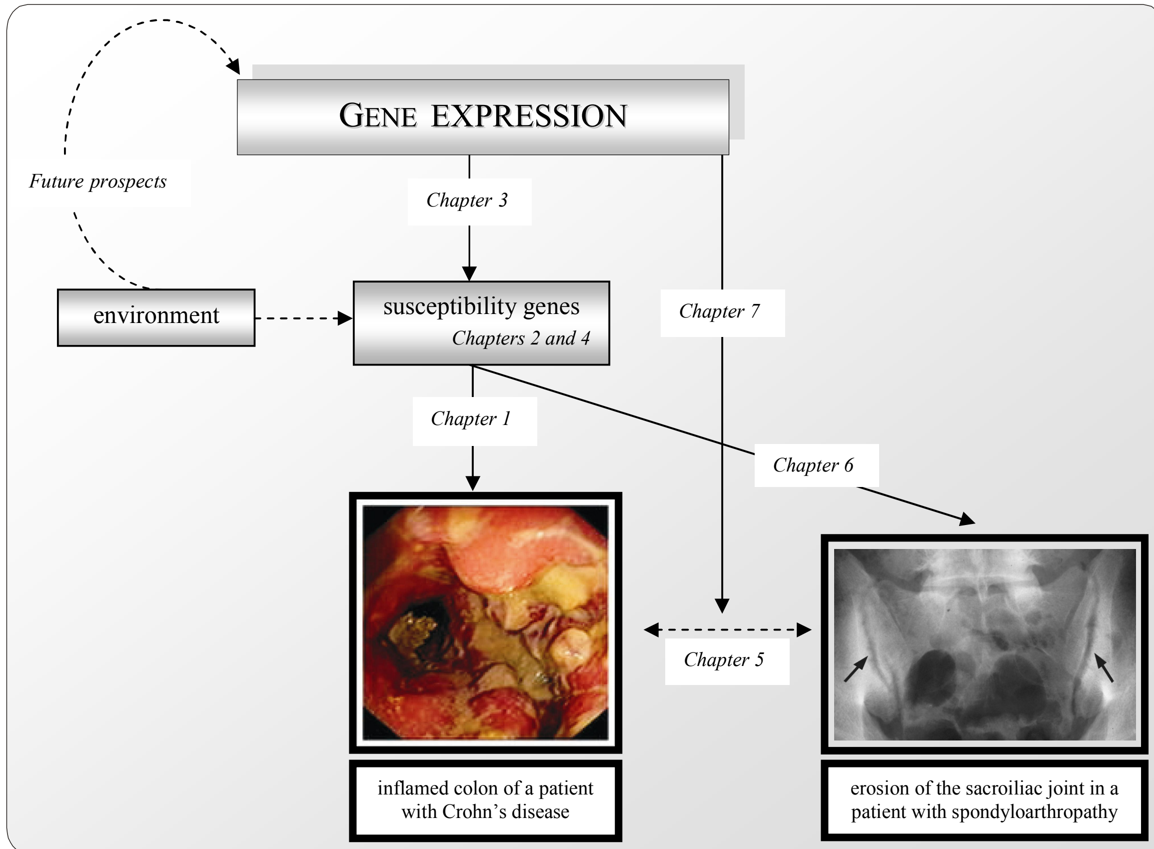
Chapter 6 113

CARD15 mutations in patients with spondyloarthropathy are linked with disease progression and evolution to Crohn's disease
Ann Rheum Dis. 2005 Jun; 64(6): 930-5

Chapter 7 127

Altered gut transcriptome in Spondyloarthropathy
Ann Rheum Dis. 2006 Feb

Future prospects 143
Summary and discussion 151
Samenvatting en discussie 159
Curriculum vitae 167
Dankwoord 171



Schematic outline of the project. *A central theme in this study was the use of gene expression analysis in the search for susceptibility genes for Crohn's disease (CD), and for the study of CD-related gut inflammation observed in spondyloarthropathy (SpA). Crohn's disease is a complex genetic trait where both genes and environmental factors play a role. In Chapter 1, the current status on genetic susceptibility in CD is reviewed. A novel functional promoter polymorphism in the recently identified CD-susceptibility gene, CARD15, is illustrated in Chapter 2. We describe a method to identify new susceptibility genes by the use of transcriptome profiling in Chapter 3, while two new candidate genes are characterized and described in Chapters 3 and 4 respectively.*

The occurrence of silent or subclinical gut inflammation in SpA patients and its relation to CD is reviewed in Chapter 5. In addition, the genetic contribution of CARD15 in SpA is studied in Chapter 6. In Chapter 7, we compared global gene expression in intestinal biopsies of SpA and CD patients, and provided additional evidence that SpA is a good model to study early CD. Finally, the influence of genetic risk factors on gene expression is described in more detail in the Future prospects, together with an allusion on the importance of epigenetic mechanisms in gene expression studies. In this last section, a new approach for the identification of susceptibility genes, based on the transmission of gene expression within families, is suggested.

Section I

The objective of this project was to use gene expression analysis for the study of genetic susceptibility in Crohn's disease (CD). Crohn's disease is a painful and incapacitating chronic inflammatory disease of the intestine. The symptoms of CD include chronic diarrhea, abdominal pain, loss of appetite, and weight loss. Approximately 1-2 in 1,000 people are affected and the incidence is still increasing. A life-long treatment is required, and the quality of life is severely influenced. Currently, there is no cure and patients suffer from a high rate of recurrence.

The cause of ongoing inflammation in the intestine of patients with CD results from a complex interaction of susceptibility genes and environmental factors. **Chapter 1** provides an overview of the current status of genetic linkage data, association studies and genetic risk factors for CD. In 2001, the first susceptibility gene for CD was cloned. This gene, called *CARD15*, is an intracellular receptor for bacterial cell-wall components, and can activate NFκB to induce an inflammatory response. Three single nucleotide polymorphisms (SNPs) within *CARD15* have been independently associated with CD, and are present in approximately 40% of CD patients. The identification of functional consequences of these SNPs is the subject of much ongoing research. Because CARD15 expression is up-regulated during inflammation, we screened its promoter for polymorphisms. In **Chapter 2**, we describe a novel SNP in the 5' untranslated region (5'UTR) of *CARD15*, located within a DNA stretch that contains a potential binding site for the E2F transcription factor. Interestingly, this SNP influences the induction of a reporter gene controlled by a part of the *CARD15* promoter in response to tumor necrosis factor alpha, which is a crucial pro-inflammatory cytokine. Furthermore, this polymorphism is linked to the known CD-associated mutations. In order to characterize this link, we are currently mapping the haplotype structure surrounding *CARD15*.

Unfortunately, the identification of new susceptibility genes for CD by classical linkage studies is complicated by genetic heterogeneity and incomplete penetrance of the disease. Therefore, in **Chapter 3**, we suggest an alternative approach for the identification of new candidate genes by the use of transcriptome analysis. This initiative was based on the idea that much of the natural variation in gene expression is genetically transmitted and probably accounts for susceptibility to common diseases such as CD. We studied genome-wide gene expression in biopsies from unaffected colon areas of CD patients and healthy controls, and focused on those genes that are differentially expressed and located near a known locus for CD. A cluster of highly related genes, encoding metallothioneins (MTs), was studied in more detail. These genes are located approximately 6 centimorgan q-telomeric from the locus containing *CARD15*. We provide evidence for MT expression as a genetic risk factor in CD, and describe a new disease-

modifying polymorphism in the main transcriptional regulator for MTs, called the MRE-binding transcription factor 1 (*MTF1*). In addition, a review of the involvement of MT in several diseases is provided in **Chapter 4**, together with current knowledge of the regulation of human MT expression.

Section II

The most common extra-intestinal manifestation of CD is inflammation of the joints and the spine. This causes pain, swelling and stiffness of the joints and the back. Spondyloarthropathy (SpA) is a group of rheumatic disorders of the spine. Interestingly, a close relationship exists between CD and SpA clinically as well as immunologically, which is discussed in **Chapter 5**. Not only is CD prevalent among patients with SpA, silent or subclinical gut inflammation has been described in up to two-thirds of patients with SpA. When *CARD15* was mapped as the first CD-associated gene, we tested the presence of the CD-associated *CARD15* mutations in a cohort of SpA patients, as described in **Chapter 6**. We provided a first genetic link between CD and SpA: the presence of CD-associated *CARD15* polymorphisms was associated with SpA patients with subclinical chronic gut inflammation. Lastly, in **Chapter 7**, we reported a gene expression survey of intestinal biopsies of SpA and CD patients, and showed that an altered gut transcriptome was found in SpA patients with subclinical chronic gut inflammation. The alterations were comparable to those seen in CD, confirming initial clinical observations and suggesting that SpA serves as a good model for early CD.

To conclude, the mounting significance of studying gene expression in genetic susceptibility studies is outlined in the **Future prospects**. We describe a new approach to identify risk genes based on the transmission of gene expression signatures in families.

-Section I-

Chapter 1

INTRODUCTION TO THE MOLECULAR GENETICS OF CROHN'S DISEASE

Debby Laukens

Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium

INTRODUCTION

Inflammatory bowel disease (IBD) is a term that refers to both Crohn's disease (CD) and ulcerative colitis (UC). Generally, the diagnosis of CD or UC is made based on the location of inflammation and depth and kind of lesions after endoscopic (Figure 1), radiologic and histologic examination. Although CD most commonly affects the terminal part of the small intestine (ileum) and the proximal part of the large intestine (colon), it may involve any section of the gastrointestinal tract. All layers of the intestine may be involved, and normal healthy bowel can exist between patches of diseased bowel. Granulomas are commonly found. In UC on the other hand, inflammation is limited to the colon. Only the superficial layers (mucosa) of the colon are affected in a symmetric and continuous distribution, starting at the level of the rectum. On the other hand, there may be a spectrum of illnesses within each disorder, suggesting distinct underlying pathogenic mechanisms. In 10 percent of the cases where only the colon is affected, it is difficult to establish a definite diagnosis, hence these are classified as indeterminate colitis (Price 1978). However, most patients with indeterminate colitis evolve to a definite diagnosis of UC or CD during follow-up. Whether UC and CD are fundamentally different diseases, or part of one concept, remains unclear and holds conceptual and practical implications. Good classification practice is essential because CD and UC differ in response to therapy and involvement of genetic and environmental risk factors (Vermeire *et al.* 2000).

Both CD and UC have a prevalence of 1-2 cases per 1,000 individuals in western countries. In Brussels, the mean annual incidence for CD has been estimated in 1992-1993 at 4.1 per 100,000 inhabitants/year (Van Gossum *et al.* 1996). The emergence of IBD in the developed nations of

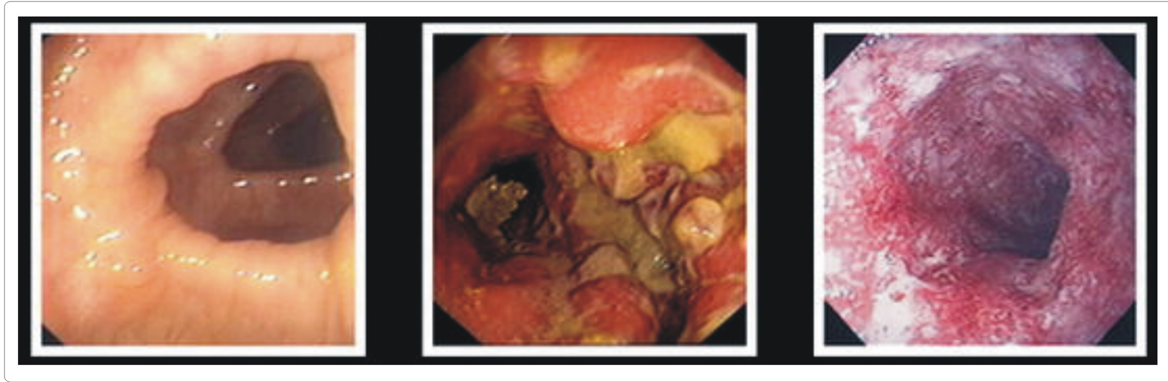


Figure 1 Colonoscopy images from healthy control, Crohn's disease and ulcerative colitis (left to right).

Western Europe and North America during the past century seems indisputable (Farrokhyar *et al.* 2001). Part of the increase in incidence reflects developments in endoscopic and other diagnostic techniques. Nevertheless, most epidemiologists accept that there was a considerable increase in CD incidence during the last century in developed countries. Western lifestyle, diet (Gibson and Shepherd 2005), exposure to infection (Hugot *et al.* 2003), and domestic hygiene (Wells and Blennerhassett 2005) are most probably involved in this process. Crohn's disease may occur among people of all ages, but it is primarily a disease of adolescents and young adults, affecting mainly those between the age of 15 and 35.

Currently, the long-term treatment of IBD is multifaceted, depending on the type of disease, and sites involved (Podolsky 2002). Most clinicians use a stepped approach to therapy, the so-called therapeutic pyramid, in which more potent agents are added to the regimen if less active drugs fail to achieve an adequate response. The primary treatment of symptoms with antidiarrheal agents itself is important. Aminosalicylates, corticosteroids and other immunomodulators are aimed at controlling inflammation. The use of antibiotics and probiotics is being investigated, especially for CD. Beneficial health effects, which might result from rebalancing of the gut flora, were observed after such treatment (Sartor 2003). Towards the top of the pyramid are the biological agents. The first biological therapy for CD, based on anti-tumor necrosis factor alpha (anti-TNF) has been approved by the Food and Drug Administration in 1998, and has turned out to be very effective (Targan *et al.* 1997). Up to 60% of CD patients need an operation within 10 years of disease onset, as a result of fistulizing, perforating, stricturing and/or obstructing complications.

The exact cause of IBD remains uncertain. It is clear that IBD arises from the interplay of environmental factors in genetically defined individuals. Today, the most accepted hypothesis about pathogenesis of IBD is that disease results from an abnormal host immune response to

bacteria that are normally found in the intestine (Sartor 2003). It is hypothesized that in normal conditions, exposure to commensal bacteria modulates the inflammatory immune response of the gut to the numerous bacteria and food antigens to which it is constantly exposed (Saxelin *et al.* 2005). In IBD, however, exposure to microflora triggers an inflammatory response by the cells lining the mucosa, leading to a chronic, destructive immune response, ultimately causing ulcerations and bowel injury. Indeed, an immune response to intestinal bacteria seems to be crucial in the pathogenesis of IBD (Sartor 2003): 1) genetically engineered rodents, such as IL10 deficient mice, develop colitis when exposed to commensal bacteria but remain disease free when raised in a sterile environment (Shi and Walker 2004); 2) the negative influence of faecal stream in the pathogenesis of recurrent CD lesions after curative resection of the distal ileum (Rutgeerts *et al.* 1991); 3) the efficacy of probiotics and antibiotics in treatment and prevention of IBD; 4) loss of tolerance to commensal bacteria (Jump and Levine 2004); 5) the increased number of surface-adherent and intracellular bacteria in the colon and ileum of IBD patients (Swidsinski *et al.* 2002; Darfeuille-Michaud *et al.* 2004); 6) the association of microbial receptor gene variants with IBD (Hugot *et al.* 2001; Klein *et al.* 2002; Franchimont *et al.* 2004); and 7) microbial DNA is found in granulomas, the most specific histological lesion found in CD (Ryan *et al.* 2004). It is, however, unclear whether the immune system is activated as a result of an intrinsic defect (either a constitutive activation or the failure of down-regulatory immune mechanisms) and/or because of continued stimulation resulting from a defective epithelial mucosal barrier.

Many researchers have invested in finding causative agents and risk factors for IBD. One of the leading infectious candidates is *Mycobacterium avium paratuberculosis* (Autschbach *et al.* 2005), but also *Listeria monocytogenes* and *Helicobacter hepaticus* have been linked to IBD. However, so far, attempts at using specific anti-mycobacterial chemotherapy have been unsuccessful (Goodgame *et al.* 2001). In any case, the complex phenotypes of IBD cannot be fully explained either by environmental factors, or by a single gene. Increasing evidence suggest a multigenic nature in combination with the above-mentioned aspects.

Crohn's disease and UC share the same ethnic predisposition, and mixed families exist in which some members are affected with CD and others with UC. However, genetic data of susceptibility genes tend to differ between CD and UC. Here, we review the progress in genetic analyses and genetic risk factors for CD.

GENETIC EPIDEMIOLOGY OF CROHN'S DISEASE

The strong influence of genetic determinants in CD has been shown by familial clustering (Peeters *et al.* 1996), and the high concordance rate, about 20-44%, in monozygotic twins (Tysk *et al.* 1988; Hugot *et al.* 1996; Thompson *et al.* 1996; Halfvarson *et al.* 2003). First-degree relatives of affected individuals show a 20 to 50-fold higher risk for developing CD (Monsen *et al.* 1991). The sibling relative risk for CD is 36.5, which is higher than the risk reported for other complex diseases such as diabetes type I and schizophrenia (Hugot *et al.* 1996; Peeters *et al.* 1996; Satsangi *et al.* 1996a). Moreover, siblings are frequently affected at similar ages, and concordance rates reach 80% for disease site, behavior and presence of extra-intestinal manifestations (Halfvarson *et al.* 2003). Ethnic differences in disease frequency have been shown. For instance, the prevalence among Ashkenazi Jews is much higher than among Sephardic Jews, even though they share a similar living environment in the same community (Fidder *et al.* 2003). An increased mother to child transmission has been observed (Akolkar *et al.* 1997), suggesting that one of the genetic risk factor is an imprinted gene, which is in agreement with the linkage of CD to the X chromosome shown in several studies (Vermeire *et al.* 2001; van Heel *et al.* 2003). All these observations are considered to be evidence for a strong genetic predisposition to the etiology of CD.

The incomplete concordance rate in monozygotic twins suggests that genetic factors are not enough to develop CD. Furthermore, the incidence of CD in immigrant populations is often different from the incidence of the population of origin, but also slightly different from the incidence in the new country (Odes *et al.* 1989). The strongest environmental risk factor for IBD is tobacco smoking. Remarkably, the risk in UC and CD is different. Smoking is harmful in both the disease progression and in the onset of CD (Cosnes *et al.* 2001), while it has a positive influence on the symptoms in UC (Beaugerie *et al.* 2001). In fact, many people who stop smoking develop UC. Risk factors influencing the disease progression include CD affecting the small intestine, age of onset above 40 years and oral contraceptive use (Logan 1999). Recently, environmental factors that might be etiologically related to CD were determined in a study based on interviewing members of affected and control families (Van Kruiningen *et al.* 2005). Besides smoking and appendicitis, dietary factors represented potential risk factors for CD. These dietary components included drinking of well water, less frequent consumption of oats, rye, and bran and more frequent eating of unpasteurized cheeses. Childhood infections, birth by caesarean section and bottle feeding have all been suggested as possible risk factors for the development for CD (Timmer 2003).

In general, two complementary approaches are used to identify determinants underlying genetically complex traits such as CD: candidate gene association studies and genetic linkage. In the candidate gene approach, functionally interesting genes are screened for polymorphisms in a limited number of patients, followed by comparing frequencies in large disease and control populations. In genetic linkage, candidate regions associated with disease are identified, scanning the entire chromosome for microsatellite markers. These so-called disease loci are usually very broad and can rarely be narrowed down to less than a few megabases. If the candidate region contains an amenable number of genes, a candidate gene is chosen by virtue of a known property, e.g. function, tissue-specific expression or structural motif, which is referred to as positional cloning. However, the mapping of CD loci is hampered by variability in phenotype, genetic heterogeneity across populations, uncontrolled environmental influences and limited statistical power. For instance, in CD, each genome scan results in a minimum of 3 loci, reflecting that more than one gene is involved in pathogenesis. The number of genes involved in the etiology of CD is unknown. Furthermore, the question remains whether common variants of a limited number of genes, each adding up small effects, account for such a common disease, or whether rare variants with a dominant effect underlie genetic susceptibility.
Although not all CD patients have a positive family history, the question arises as to the precise relationship between sporadic and familial cases of CD. Will genetic variants identified in familial cases also be pathogenic in sporadic CD patients? In a recent study addressing this question, no difference in frequency of three CD-associated variants within the *CARD15* gene (see below) was found between familial and non-familial cases (Vermeire *et al.* 2002). This suggested that these variants are partly causative for idiopathic CD patients, and are also transmitted within CD families (Esters *et al.* 2004).

GENOME-WIDE LINKAGE STUDIES IN IBD

At present, 13 genome-wide linkage studies have been performed in IBD (Hugot *et al.* 1996; Satsangi *et al.* 1996b; Cho *et al.* 1998; Hampe *et al.* 1999; Ma *et al.* 1999; Duerr *et al.* 2000; Rioux *et al.* 2000; Cavanaugh 2001; Paavola-Sakki *et al.* 2003; Shaw *et al.* 2003; van Heel *et al.* 2003; Barmada *et al.* 2004; Vermeire *et al.* 2004). More than 20 loci have been identified, and many of them have been independently replicated (*IBD1-9*, 3q, 4q, 6q, 7q, 8q, 10p, 22q, Xp). Table 1 covers all original articles on linkage analysis performed on IBD or mixed families from 1996 until now. Data on UC only were not included. These analyses uncover great ethnic differences as to loci and degree of linkage. Identification of the precise genes within these loci

Table 1 Susceptibility loci for Crohn's disease

IBD Locus	Chromosomal position	Marker at highest score	Score	Population	Reference
IBD7	1p13.1	D1S252	NPL 2.61	Flemish IBD	(Vermeire *et al.* 2004)
IBD7	1p33-32.3	D1S197	NPL 2.07	Flemish IBD	(Vermeire *et al.* 2004)
IBD7	1p36.13-36.11	D1S552-D1S234	LOD 2.64	Non-Jewish American	(Cho *et al.* 1998)
	1q21.3	D1S305	NPL 2.97	Flemish IBD	(Vermeire *et al.* 2004)
	1q43-q44	D1S2670-D1S2682	LOD 2.08	North European IBD	(Hampe *et al.* 1999)
	2q32.3	D2S117	LOD 1.25	American IBD	(Duerr *et al.* 2000)
IBD9	3p14.2-14.1	D3S1766-D3S1285	LOD 2.4	Canadian IBD	(Rioux *et al.* 2000)
IBD9	3p21.31	D3S1573	LOD 2.69	North European IBD	(Satsangi *et al.* 1996b)
	3q13.12	D3S3045	LOD 1.31	American IBD	(Duerr *et al.* 2000)
	3q25.1-26.31	D3S1279-D3S3725	LOD 2.1	North European CD	(van Heel *et al.* 2003)
	3q26.31	D3S3053-D3S2427	LOD 3.23	American IBD	(Cho *et al.* 1998)
	4q25	D4S406	NPL 2.07	Finnish IBD	(Paavola-Sakki *et al.* 2003)
	4q25	D4S406	NPL 2.17	Flemish IBD	(Vermeire *et al.* 2004)
	4q25	D4S2623	LOD 3.0	American IBD	(Cho *et al.* 1998)
	5q14.1	D5S1501	LOD 1.69	Caucasian CD	(Barmada *et al.* 2004)
IBD5	5q33-35	-	LOD 2.2	Jewish American CD	(Ma *et al.* 1999)
IBD3	6p	D6S197	LOD 3.06	North European CD	(Shaw *et al.* 2003)
IBD3	6p22.2	D6S2439	LOD 2.6	Caucasian CD	(Barmada *et al.* 2004)
IBD3	6p22.2-21.2	D6S1281-D6S1019	LOD 2.3	Canadian IBD	(Rioux *et al.* 2000)
IBD3	6p23-22.3-22.2	D6S289-D6S276	LOD 2.07	North European IBD	(Hampe *et al.* 1999)
	6q24.1	D6S314	NPL 2.44	Flemish IBD	(Vermeire *et al.* 2004)
	6q25.2-26	D6S2436-D6S305	LOD 2.21	Caucasian IBD	(Barmada *et al.* 2004)
	7q21.11	D7S669	LOD 3.08	North European IBD	(Satsangi *et al.* 1996)
	7q32.1-31.33	D7S40-D7S648	LOD 0.91	American IBD	(Duerr *et al.* 2000)
	8q12.1-13.1	D8S1113-D8S1136	LOD 1.57	Caucasian IBD	(Barmada *et al.* 2004)
IBD2	12q21.1-21.2	D12S303-D12S326	LOD 1.82	North European CD	(Hampe *et al.* 1999)
IBD2	12q23.3	D12S78	LOD 2.34	Finnish CD	(Paavola-Sakki *et al.* 2003)
IBD4	14q11.2	-	LOD 8.2	American CD	(Ma *et al.* 1999)
IBD4	14q11.2	D14S261	LOD 3.6	American CD	(Duerr *et al.* 2000)
IBD4	14q11.2	D14S50	LOD 2.34	Flemish IBD	(Vermeire *et al.* 2004)
IBD4	14q12	D14S80	LOD 2.04	Flemish IBD	(Vermeire *et al.* 2004)
IBD4	14q13.1	D14S49	LOD 2.44	Flemish IBD	(Vermeire *et al.* 2004)
	15q26.1	D15S652	LOD 1.82	Caucasian IBD	(Vermeire *et al.* 2004)
	15q26.1-26.2	D15S652-D15S816	LOD 2.02	Caucasian IBD	(Barmada *et al.* 2004)
IBD8	16p13.13	D16S748	NPL 2.49	American CD	(Cho *et al.* 1998)
IBD1	16q12.1	D16S411-D16S419	LOD 5.79	multicentre (IBDIGC)	(Cavanaugh *et al.* 2001)
IBD1	16q12.1	D16S409-D16S411	LOD 1.71	European CD	(Hampe *et al.* 1999)
IBD1	16q12.1-12.2	D16S409, D16S419	-	Caucasian CD	(Hugot *et al.* 1996)
	16q21-23.1	D16S514-D16S515	LOD 2.2	North European CD	(Van Heel *et al.* 2003)
	18p11.31	D18S62	NPL 2.3	Flemish IBD	(Vermeire *et al.* 2004)
	18q22.2	D18S61	LOD 1.15	American IBD	(Duerr *et al.* 2000)
IBD6	19p13.3	D19S591-GATA21G05	LOD 4.6	Canadian CD	(Rioux *et al.* 2000)
IBD6	19p13.3-13.2	D19S1034-D19S586	LOD 1.38	American IBD	(Cho *et al.* 1998)
	19q13.31	D19S217	LOD 2.9	North European CD	(Van Heel *et al.* 2003)
	20p12.3	D20S192	NPL 2.6	Flemish IBD	(Vermeire *et al.* 2004)
	22q11.23-12.1	D22S315-D22S421	LOD 1.52	North European IBD	(Hampe et al. 1999)
	22q12.1	D22S689	LOD 1.5	Caucasian IBD	(Barmada et al. 2004)
	22q13.31	D22S274	NPL 1.19	Flemish IBD	(Vermeire et al. 2004)
	Xp21.2-21.3	DXS1202-DXS1214	LOD 2.0	North European IBD	(Hampe et al. 1999)
	Xp22.11-21.2	DXS1226-DXS1214	LOD 1.59	North European CD	(Van Heel et al. 2003)
	Xq28	DXS99	LOD 1.7	Flemish IBD	(Vermeire et al. 2004)

LOD: logarithm of the odds; NPL: non-parametric LOD score

is now a crucial step in understanding the contribution and interaction of different gene products in disease onset and pathology. Typically, genome-wide scans identify large loci, spanning several hundreds of genes. Fine mapping of these loci is subsequently used to finally end up with an amenable amount of genes for further study. The latter approach has been successfully employed for the *IBD1* locus (16q12), which has resulted in the identification of the first major susceptibility gene for CD, *CARD15* (see next paragraph).

The ultimate goal of characterizing new susceptibility genes is a more accurate prediction of disease, which might permit the use of more specific therapy adapted to an individual's genotype. Furthermore, the identification of these variants offers the possibility to stratify genome-wide linkage studies, increasing the power of identification of new loci (Shaw *et al.* 2003; van Heel *et al.* 2003).

SUSCEPTIBILITY GENES AND DETERMINANTS FOR CD

The majority of early studies of genetic susceptibility in CD used case-control association studies, in which allele frequencies of candidate genes are compared between patients and healthy controls. However, such studies have been criticized, basically because of the lack of reproducibility: original association studies often indicate a much stronger effect than subsequent ones. The reasons for failing to replicate disease associations have been discussed elsewhere (Colhoun *et al.* 2003). On the other hand, positional cloning has led to the identification of major genetic variants associated with CD on chromosome 16 (*IBD1* locus, *CARD15*) (Hugot *et al.* 2001), chromosome 10 (*DLG5*) (Stoll *et al.* 2004) and chromosome 5 (*IBD5* locus, 5q31 risk haplotype) (Rioux *et al.* 2001). These genes will be discussed below.

CARD15

The *IBD1* locus represents the best replicated region, showing linkage to CD and specifically not to UC. It was originally mapped in 1996 (Hugot *et al.* 1996), and has since been replicated in many studies (Table 1) (Ohmen *et al.* 1996; Cho *et al.* 1997; Brant *et al.* 1998; Cavanaugh *et al.* 1998; Curran *et al.* 1998; Annese *et al.* 1999; Hampe *et al.* 1999; Akolkar *et al.* 2001). Furthermore, an international collaborative study on IBD reported a remarkable high linkage score for CD at position D16S411 (Cavanaugh 2001). In 2001, two groups simultaneously identified *NOD2* (nucleotide oligomerization domain 2), now officially called *CARD15* (caspase activation and recruitment domain 15), as the first susceptibility gene for CD (Hugot *et al.* 2001; Ogura *et al.* 2001a). Hugot and colleagues employed the positional cloning strategy while Ogura

and co-workers identified *CARD15* by the positional candidate gene approach. European and North American patients with CD, including those without a family history, are more likely to have variants in *CARD15* than those without CD. It must, however, be noted that another gene within *IBD1* might additionally be responsible for the linkage with this region. This became even more clear when a genome-wide linkage study was stratified with the *CARD15* variants, and linkage to chromosome 16 was observed in CD patients ~25 centimorgan q telomeric of *CARD15* (van Heel *et al.* 2003).
A recent study comparing a European and Korean cohort for the common *CARD15* variants illustrated that these variants do not exist in the Korean cohort. Moreover, in this study, no association was found between other SNPs in *CARD15* and CD (Croucher *et al.* 2003). This implicates that the CD-associated polymorphisms arose after the European/Asian separation. Similarly, no *CARD15* mutations were found in the Chinese (Leong *et al.* 2003) and Japanese (Inoue *et al.* 2002) population, and the association of *CARD15* variants is not as strong among the Finns (Helio *et al.* 2003), Irish (Bairead *et al.* 2003), New Zealands (Leung *et al.* 2005) and the Scottish (Arnott *et al.* 2004). Hence, although ethnically divergent populations may present identical phenotypes, they do not necessarily share the same set of predisposing genes.

Polymorphisms in CARD15 associated with CD. The CARD15 protein belongs to the NOD1/Apaf-1 family, which comprises cytosolic proteins that are composed of an N-terminal caspase recruitment domain (CARD), a centrally located nucleotide-binding domain (NBD), and a C-terminal leucine-rich regulatory (LRR) domain (Bertin *et al.* 1999; Inohara *et al.* 1999; Ogura *et al.* 2001b). Three common single nucleotide polymorphisms (SNPs) in *CARD15* were independently associated with CD (Figure 2): two missense mutations R702W (c.2104C>T, SNP8) and G908R (c.2722G>C, SNP12), and one frameshift mutation 1007fs (c.3020insC, SNP13), which truncates the protein with 30 amino acids (Hugot *et al.* 2001). All three variants alter the C-terminal third of the protein, and are within or close to the LRR domain, which is involved in ligand recognition. The heterozygous carrier frequency of these variants within CD ranges from 30 to 50%, while 3 to 15% is homozygous or compound heterozygote, meaning that they carry two variants (Hugot *et al.* 2001; Ogura *et al.* 2001a; Abreu *et al.* 2002; Ahmad *et al.* 2002; Hampe *et al.* 2002). By comparison, 8 to 15% of healthy controls are heterozygous and 0 to 1% carries a homozygous variant. The relative risk of developing CD if someone carries one of these variants is increased by a factor of 1.5 to 3 for heterozygous carriers, but the risk increases with a factor of 20 to 40 if someone is homozygous or compound heterozygous. Although this relative risk seems high, it must be stated that the absolute risk for developing CD

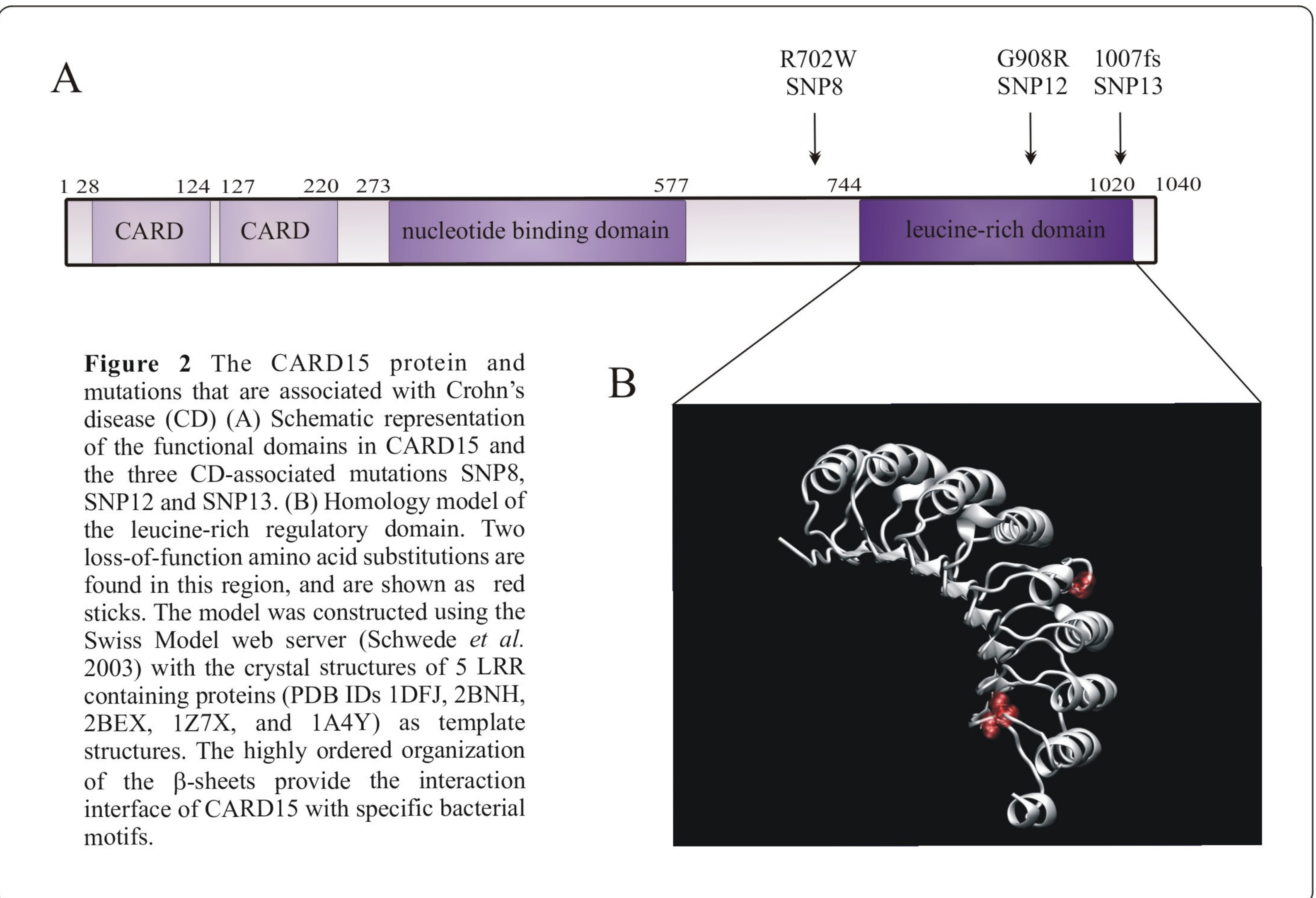


Figure 2 The CARD15 protein and mutations that are associated with Crohn's disease (CD) (A) Schematic representation of the functional domains in CARD15 and the three CD-associated mutations SNP8, SNP12 and SNP13. (B) Homology model of the leucine-rich regulatory domain. Two loss-of-function amino acid substitutions are found in this region, and are shown as red sticks. The model was constructed using the Swiss Model web server (Schwede *et al.* 2003) with the crystal structures of 5 LRR containing proteins (PDB IDs 1DFJ, 2BNH, 2BEX, 1Z7X, and 1A4Y) as template structures. The highly ordered organization of the β-sheets provide the interaction interface of CARD15 with specific bacterial motifs.

is no more than 1 in 25 for homozygous carriers. This reduced penetrance can undoubtedly be explained by the requisite of environmental risk factors and/or additional genetic determinants.
A number of recent reports indicate that common genetic variation in the human genome exists as discrete haplotype blocks, each with a limited diversity. The haplotype structure in a 177-kb region surrounding the *CARD15* gene has been determined (Vermeire *et al.* 2002). A specific haplotype was association with CD, and the three CD-associated SNPs were all unique subvariants of this haplotype.
Different mutations in *CARD15* have been associated with increased susceptibility to psoriatic arthritis (Rahman *et al.* 2003), and to Blau syndrome (Miceli-Richard *et al.* 2001), a rare autosomal dominant disorder characterized by early-onset granulomatous arthritis, uveitis, and skin rash with camptodactyly. Interestingly, the mutations found in Blau syndrome are located in the NBD, suggesting a different molecular pathogenesis between Blau syndrome and CD.

Functional studies on CARD15 and implications for CD. The Nod1/Apaf-1 family of proteins displays striking similarity to a class of disease resistance (R) proteins found in plants. Following specific recognition of pathogen products, these R proteins mediate a defence response associated with metabolic alterations and localized cell death at the site of pathogen invasion. The LRR domains of R proteins are highly diverse and appear to be involved in the recognition of a wide array of pathogen components (Dixon *et al.* 2000). Similar to the R proteins, CARD15 appears to play an important role in innate and acquired immunity as a sensor of bacterial components. Specifically, CARD15 participates in the signalling events triggered by host recognition of specific bacterial motifs, and subsequently activates NFκB, the key mediator in the production of pro-inflammatory mediators (Inohara and Nunez 2003). Naturally occurring peptidoglycan (PG) fragments were identified as the microbial motifs sensed by CARD15, more specifically muramyl dipeptide (MDP, GlcNAc-MurNAc), found in Gram-negative and Gram-positive bacterial PGs (Girardin *et al.* 2003). However, direct binding of MDP to CARD15 has not yet been demonstrated.
The expression of CARD15 was first thought to be restricted to cells of the myeloid lineage, primarily to monocytes (Ogura *et al.* 2001b; Gutierrez *et al.* 2002). With more accurate techniques, however, expression of CARD15 was shown in epithelial cells (Hisamatsu *et al.* 2003; Rosenstiel *et al.* 2003), keratinocytes (Voss *et al.* 2005), vascular endothelial cells (Oh *et al.* 2005) and paneth cells (Lala *et al.* 2003; Ogura *et al.* 2003). Moreover, its expression is enhanced by pro-inflammatory cytokines and bacterial components via NFκB, a mechanism that may contribute to the amplification of the innate immune response (Gutierrez *et al.* 2002;

Rosenstiel *et al.* 2003). Consistent with this observation, an elevated expression of CARD15 has been shown in inflamed areas of colonic tissue of CD patients (Berrebi *et al.* 2003).

Although the function of CARD15 in bacterial sensing is widely accepted, its physiological function is less well understood. Consequently, the implications of the CD-associated mutations in disease onset and progression remain unclear. Taken together, five hypotheses have been postulated, and they involve both loss-of-function and gain-of-function of the *CARD15* mutations, although the gain-of-function hypotheses have been criticized (Figure 3). The gain-of-function hypotheses evolved from *CARD15* knockout and transgenic mice, but are not consistent with what is seen in humans. It is important to note that these hypotheses are not mutually exclusive, and may be physiologically relevant in combination.

In vitro and *ex vivo* experiments indicated that the three CD-associated polymorphisms actually decreased the activation of NFκB and pro-inflammatory cytokine production, which is inconsistent with the observation that NFκB is up-regulated in patients (Ogura *et al.* 2001b; Bonen *et al.* 2003; Hisamatsu *et al.* 2003). However, this might reflect a lack of triggering a primary innate immune response to bacterial invasion. Since the three CD-associated variants of *CARD15* are located in or near the LRR (Figure 2), it was suggested that bacterial sensing is impaired, thus explaining the susceptibility to disease. As a consequence, clearing of bacterial products is inefficient, which might lead to a secondary, compensatory activation of NFκB independent from CARD15 (Figure 3A). Notably, peripheral blood mononuclear cells (PBMC) from individuals homozygous for the major disease-associated SNP13 mutation did not respond to synthetic MDP (Chamaillard *et al.* 2003; Inohara *et al.* 2003; Inohara *et al.* 2005). They also show a defective pro-inflammatory cytokine release after stimulation of mononuclear cells with MDP (Li *et al.* 2004; Netea *et al.* 2004; Netea *et al.* 2005). We have recently shown that PBMC isolated from CD patients carrying *CARD15* polymorphisms, produced significantly less IL1β, IL6 and IL10 after stimulation with *adherent-invasive E. coli LF82* in a gene-dose effect (Peeters *et al.* 2006). This was the first study where aberrations were found in heterozygous carriers of SNP8 and SNP12 *CARD15* mutations, which is the largest group of patients. Moreover, *CARD15* 1007fs carriage was also associated in a gene-dose-dependent manner with low mononuclear cell TNF release by stimulation with a combination of interferon gamma (IFNG) and GM-CSF (Halme *et al.* 2004).

The finding that CARD15 is expressed in paneth cells is particularly of interest, because these cells provide the host defence against microbes in the ileum, while they are not present in the normal colon. Paneth cells secrete a number of antibacterial substances in the lumen of the crypt,

like lysozyme, phospholipase A and α- and β-defensins (Ouellette and Bevins 2001). Lysosyme is an enzyme that breaks down bacterial cell wall components into MDP, which is in turn recognized by CARD15. Moreover, the ileum is an important pathological site in CD, and *CARD15* mutations are primarily present in patients with ileal involvement (Cuthbert *et al.* 2002). Ileal expression of the α-defensins was diminished in active regions in patients with *CARD15* mutations, while this was not the case in the diseased colon (Wehkamp *et al.* 2004). CARD15 is also an inducer of *β*-defensins, an effect that is lost in *CARD15* SNP13 homozygous patients, leading to a defective epithelial defence, proliferation of bacteria and potential loss of epithelial barrier function (Figure 3B).

In human monocyte-derived dendritic cell cultures, CARD15 agonists synergistically induce IL12 production in combination with Toll-like receptor 3 (TLR3), TLR4, and TLR9 agonists to induce T helper type 1 (Th1)-lineage immune responses (Tada *et al.* 2005). This synergistic effect was lost in patients carrying a mutant CARD15 protein (van Heel *et al.* 2005b; Kramer *et al.* 2006). The inflammatory phenotype of CD is difficult to reconcile with a decrease in TNF and IL12 (Sartor 1994). However, it was recently demonstrated that the synergistic induction of IL10 in response to MDP and TLR-stimuli was lost in *CARD15* mutant monocyte-deriven dendritic cells (Kramer *et al.* 2006). Interleukin 10 is crucially involved in the down-regulation of the inflammatory process. It was thus postulated that the IL10 mediated immune suppression is impaired and the counter-effect for pro-inflammatory cytokines is lost, thus contributing to chronic ongoing inflammation in CD (Figure 3C).

Incubation of normal murine macrophages with MDP was shown to suppress IL12 secretion induced by stimulation with TLR2 ligands (Watanabe *et al.* 2004). This suppression did not occur in cells lacking CARD15 or cells expressing a mutant form of CARD15 in transfection experiments. Once secreted, IL12 promotes IFNG and growth and differentiation of Th1 cells (Figure 3D). A major concern however is the reproducibility of these results (Kobayashi *et al.* 2005). Furthermore, CARD15 and TLR2 stimulation of human PBMC isolated from patients with the 1007fs mutation, led to a loss of synergistic induction of pro-inflammatory cytokines, which is also inconsistent with a TLR2 inhibitory function of CARD15 (van Heel *et al.* 2005a).

Macrophages of mice carrying a mutant *CARD15* equivalent to the 1007fs mutation, were shown to secrete higher levels of the mature form of IL1β upon stimulation with MDP (Maeda *et al.* 2005). This suggests that the variant CARD15 protein promotes processing of proIL1β to the mature IL1β. Given that IL1β-converting enzyme (ICE) is critical for this processing, it is plausible to assume that mutant CARD15 can activate ICE in response to MDP (Figure 3E).

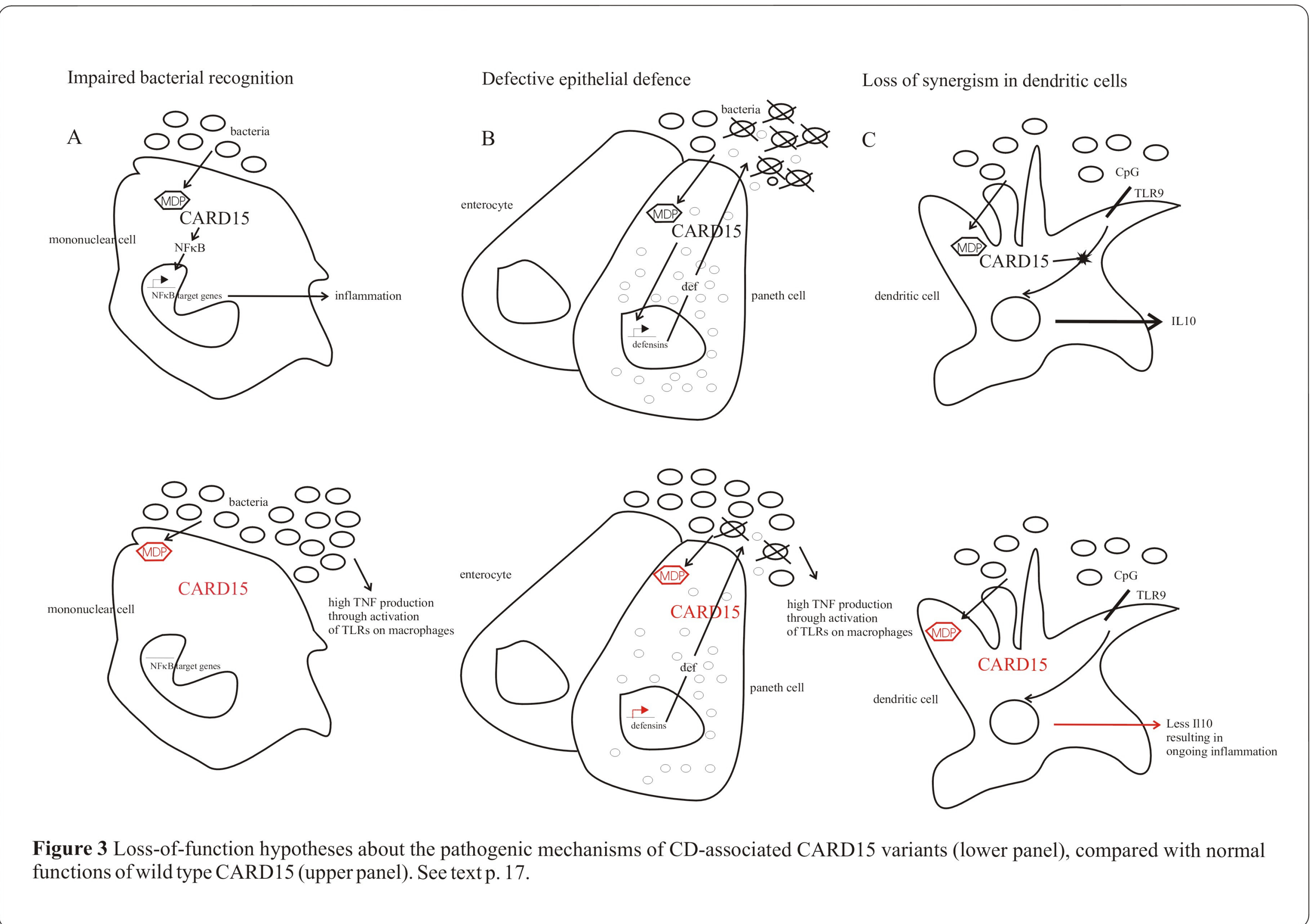


Figure 3 Loss-of-function hypotheses about the pathogenic mechanisms of CD-associated CARD15 variants (lower panel), compared with normal functions of wild type CARD15 (upper panel). See text p. 17.

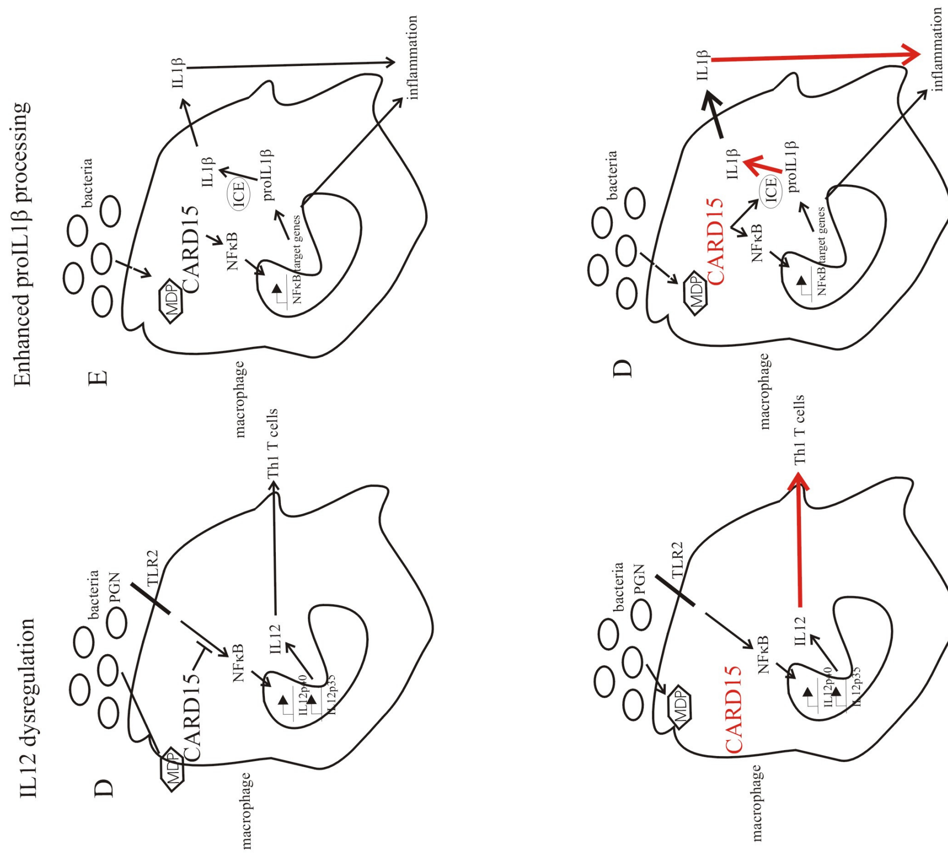


Figure 3 Gain-of-function hypotheses about the pathogenic mechanisms of CD-associated CARD15 variants (lower panel), compared with normal functions of wild type CARD15 (upper panel). See text p. 18.

However, macrophages from patients with mutant *CARD15* were shown to have reduced IL1β release upon MDP stimulation (Li *et al.* 2004). The last two observations may indicate important differences in the function of CARD15 in murine versus human cells.

Recently, the pathway for CARD15 signalling has been studied. Upon binding of MDP, CARD15 interacts with RIPK2 (receptor-interacting serine-threonine kinase 2), a kinase that phosphorylates the inhibitor of NFκB kinase, through homophilic CARD-CARD interaction, leading to the nuclear translocation of NFκB (Kobayashi *et al.* 2002). CARD15 activation also leads to ubiquitinylation of the NFκB essential modulator (NEMO) (Abbott *et al.* 2004), and has also been shown to interact with erbb2 interacting protein (ERBIN) (Chen *et al.* 2004) and GRIM-19, a protein with homology to the NADPH dehydrogenase complex (Barnich *et al.* 2005b), both essential for NFκB activation. It was shown that membrane targeting, which is crucial for NFκB activation is impaired in the presence of a SNP13 mutation (Barnich *et al.* 2005a). Two leucine residues and a tryptophan-containing motif in the C-terminal domain of CARD15 mediate this membrane interaction.

The association of *CARD15* with CD may provide a unifying explanation for several factors influencing the development of this disease. Firstly, CARD15 responds to bacterial cell wall components, linking intra-luminal bacteria to CD. Secondly, mutations of the LRR domain affect its sensing function leading to aberrant activation of the NFκB pathway, which is abnormally activated in CD (Schreiber *et al.* 1998; Tanabe *et al.* 2004). Thirdly, several studies have indicated that mutations of the *CARD15* gene are associated with ileal involvement (Ahmad *et al.* 2002; Cuthbert *et al.* 2002; Lesage *et al.* 2002), where CARD15 is highly expressed. Thus, *CARD15* seems to be important for the pathogenesis of CD.

Together with association of CD with *TLR4* and *CD14* (see p.23-25), the involvement of *CARD15* demonstrates the importance of the innate immune response in the pathogenesis of CD, which was discussed in the introduction.

DLG5

IBD-associated variants responsible for the linkage observed at 10q23 (Hampe *et al.* 1999) were attributed to the disk large gene 5 (*DLG5*) by positional cloning (Stoll *et al.* 2004).

Association with the IBD phenotype was strongest, and it was higher in the CD subgroup as compared to the UC subgroup, although this could be ascribed to the smaller UC sample size in this study. This association has since been replicated in three studies (Daly *et al.* 2005; Newman *et al.* 2006; Tenesa *et al.* 2006), while two other studies could not confirm these results (Noble *et*

al. 2005b; Torok *et al.* 2005). Two distinct haplotypes were identified with a distortion of transmission in trios. One of the risk-associated *DLG5* haplotypes was distinguished from the common haplotype by a nonsynonymous SNP c.113G>A, resulting in the amino acid substitution R30Q. This SNP was significantly associated with IBD, as was another nonsynonymous SNP c.4136C>A resulting in a P1371Q substitution. Furthermore, a significant difference in association of the 113A variant with CD was observed in affected individuals carrying the risk associated *CARD15* alleles. This suggests a complex pattern of gene-gene interaction between *DLG5* and *CARD15*.

DLG5 was first cloned in 1998 as a homolog of the drosophila DLG by searching expressed sequence tag databases for related sequences (Nakamura *et al.* 1998). It encodes a scaffolding protein essential for the maintenance of epithelial integrity (Humbert *et al.* 2003). DLG5 co-localizes to vinexin, involved in cytoskeletal organization and signal transduction, and to β-catenin, a major adherens junction protein (Wakabayashi *et al.* 2003). The expression of DLG5 has been shown in the intestine and isolated epithelial cells, and thus it seems plausible that mutations in this gene might disrupt epithelial barrier function of the intestine (Stoll *et al.* 2004). *In silico* analysis of the R30Q and P1371Q suggested that both variants probably impair scaffolding functions of DLG5.

OCTN1 and OCTN2

A locus of approximately 250 kb at 5q31 (*IBD5*) was previously associated with CD (Ma *et al.* 1999; Rioux *et al.* 2000; Giallourakis *et al.* 2003; Negoro *et al.* 2003). It was an interesting region because it harbours a cytokine cluster (Rioux *et al.* 2001). Peltekova and co-workers re-sequenced five genes included within this region, and reported two variants in the organic cation transporter gene cluster (*OCTN*): a missense substitution in *SLC22A4*, coding for OCTN1 (c.1672C>T, L503F) and a G>C transversion in the promoter of *SLC22A5*, coding for OCTN2 (c.-207G>C) (Peltekova *et al.* 2004). They are both part of a haplotype associated with CD. Moreover, the risk for disease was greater in the presence of both the *IBD5* haplotype and the CD-associated *CARD15* alleles. These observations have since been repeated twice, and genotype-phenotype analysis revealed an association particularly with colonic disease (Torok *et al.* 2005; Waller *et al.* 2005). However, a study from Noble and co-workers could not find an association between the *OCTN1/2* variants and CD in the absence of the *IBD5* haplotype (Noble *et al.* 2005a), thus a causative role for these genes remains plausible but is not yet proven. Further genetic and functional data are required to fully designate these genes as susceptibility genes. In a recent study, an association was found between the promoter polymorphism in

SLC22A5 and psoriatic arthritis (Ho *et al.* 2005), another chronic inflammatory disease that previously has been associated with *CARD15* (Rahman *et al.* 2003).
OCTN1 and OCTN2 are transmembrane proteins that regulate carnitine and cation transport (Burckhardt and Wolff 2000). They are expressed in cell types that are affected in CD, including epithelial cells, macrophages and T cells (Peltekova *et al.* 2004). The L503F variant is located within the transport domain, and thus alters its transport function: carnitine uptake in fibroblasts was lower in the L503F variant, while tetraethyl ammonium uptake was increased. The promoter polymorphism in *SLC22A5* impaired the binding of the heat shock transcription factor 1. Consequently, the expression of c.-207C OCTN2 was less induced in response to heat shock. These observations could be consistent with an altered uptake of physiological compounds or bacterial metabolites.

Other gene associations for CD

Many genes have been tested for association to CD or IBD in general, because they are positional or functional candidate genes (Table 2). We describe three genes that have been replicated in association studies (*TLR4, CD14* and *TNF*), because of their regulatory role in innate immune responses and functional relevance in IBD pathogenesis.

TLR4. Lipopolysaccharide (LPS), a cell wall component of Gram-negative bacteria, is a major inducer of inflammation, and its signalling is mediated through the cell surface TLR4. During intestinal inflammation, TLR4 is up-regulated on epithelial cells, macrophages and dendritic cells, thus providing a first line defence against enteric Gram-negative bacteria. An association between a polymorphism in the LRR region of TLR4 (c.896A>G, D299G) has been reported within a Dutch CD and UC cohort (Franchimont *et al.* 2004). Allele frequencies of 10.9% were found in CD patients, versus 5% in healthy controls. The association was replicated twice (Braat *et al.* 2005; Gazouli *et al.* 2005), but could not be reproduced in three other studies (Arnott *et al.* 2004; Torok *et al.* 2004; Lakatos *et al.* 2005). This mutation was previously linked to a decreased bronchial responsiveness to LPS (Arbour *et al.* 2000), and impairs LPS signaling. However, no functional defect, e.g. cytokine release, LPS recognition, has been attributed to heterozygous carriership of this mutant in CD patients (Erridge *et al.* 2003; von Aulock *et al.* 2003). A transcriptome analysis of monocyte-deriven dendritic cells, isolated from homozygous *CARD15* and *TLR4* patients, revealed that a large number of genes are differentially regulated in both groups (Braat *et al.* 2005). This suggests that mutations in different genes can cause similar effects on gene transcription and can thus result in a similar phenotype.

Table 2 Positional and functional candidate genes for CD

	Locus	Chromosome	Candidate genes		
			association	**no association**	**not done**
positional	IBD1	16q	CARD15		CDH1
	IBD2	12q	KRTHB6	SLC11A2, STAT6, AVIL, ITGB7, IFNG, VDR	MMP19
	IBD3	6p	TNF, HLA, HSPA1B, MLN, TAP2	BMP6, PSORS1C1, F13A1, MICA, MICB, TAP1, LTA	PSMB9, PSMB8, C4A, C2
	IBD4	14q	NFKBIA		TCRA, TCRD, proteasome cluster, LTB4R2
	IBD5	5q	CD14, OCTN1, OCTN2, IL4	IL12B	IL3, IL5, IL13, CSF2
	IBD6	19p	ICAM1, C3		TBXA2R, CYP4F3, MADCAM1
	IBD7	1p	PTPN22	GSTM1, MTF1, MASP2	HSPG2, CASP9
	IBD8	16p		IL4R, ITGAM, IGSF6, CD19	SPN, ITGAL
	IBD9	3p	TLR9, MLH1	CCR2, CCR5, GNAI2	IL5RA
		1q	EPHX1	SELE, SELL	TGFB2, LEFTY2
		4q			IL2
		7p	CARD4	EGFR, IL6	MUC17
		7q	ABCB1, SERPINE1		MET
		10p			ITGB1
		10q	DLG5		
		17q			ICAM2
		X			
functional		1q	IL10		
		2q	IL1, IL1R1, CTLA4, SLC11A1	UGT1A7	
		4p	TLR1, TLR6		
		4q	TLR2		
		8p		NAT2	
		9q	TLR4		
		11q	IL18, MMP3	GSTP1	
		12p	TNFRSF1A		
		15q	IL16	CYP1A1	
		17q	CCL2		
		20q	BPI		
		21q		AIRE	TFF1, TFF2, TFF3
		22q		GSTT1	

CD14. The *CD14* gene is located within the *IBD5* locus, and encodes two protein forms: a glycosylphosphatidylinositol-anchored membrane protein and a monocyte- or liver-derived soluble serum protein that lacks the anchor. Both molecules are critical for LPS dependent signal transduction (Wright *et al.* 1990). Klein and co-workers described an increased incidence of *CD14* c.-159C>T heterozygous and homozygous mutants in CD patients compared to healthy controls (Klein *et al.* 2002). This association was replicated in a Greek population (Gazouli *et al.* 2005), but it was not observed in three other studies (Arnott *et al.* 2004; Torok *et al.* 2004; Peters *et al.* 2005). An interaction between *CARD15* and *CD14* has been observed (Klein *et al.* 2003). Although the expression of CD14 on monocytes is higher in CD as compared to healthy controls, its expression did not correlate to the *CD14* genotype (Griga *et al.* 2005). Similarly, CD14 expression is higher in inflamed intestinal mucosa (Rogler *et al.* 1997), but the association to the *CD14* genotype was not found in this study.

TNF. Tumor necrosis factor alpha (TNF) is a pro-inflammatory cytokine that provides a rapid form of host defence against infection but is fatal in excess. It was a strong candidate gene within *IBD3*, because TNF levels are increased in the serum, mucosa and stool of IBD patients, while anti-TNF therapy is very efficacious in IBD (Targan *et al.* 1997; Present *et al.* 1999). TNF production is under a strong genetic influence (Westendorp *et al.* 1997). Three SNPs in the promoter of *TNF* (c.-1031C, c.-863A and c.-857T) were associated with susceptibility to CD in a Japanese population (Negoro *et al.* 1999). In contrast, none of these SNPs could be associated with IBD in two independent Caucasian populations, while the c.-857C allele was more prevalent in IBD and UC (van Heel *et al.* 2002). Interestingly, this variant was also associated with CD when they left out the common *CARD15* allele carriers, meaning that these genes act independently to confer CD susceptibility. *Ex vivo* LPS challenging of monocytes resulted in a higher production of TNF in c.-857C homozygous individuals. In addition, the transcription factor OCT1, an inhibitor of NFκB transactivation, specifically binds to the c.-857T variant, and not to the c.-857C.

Pediatric onset, colonic disease and familial aggregation of CD was associated with a frequent polymorphism in the binding site for NFκB in the *TNF* promoter, c.-863C>A (Levine *et al.* 2005). Electrophoretic mobility shift assay and transfection experiments demonstrated that MDP exposure stimulates *TNF* gene transcription, as a result of CARD15-induced NFκB activation (Linderson *et al.* 2005). When the CD-associated *CARD15* 1007fs variant was analyzed, induction of *TNF* promoter activity was found to be defective. Different combinations

of *CARD15* and *TNF* promoter polymorphisms gave rise to distinct TNF transcription levels, which means that *CARD15* and *TNF* promoter polymorphisms interact to exert a functional effect on MDP induced TNF production. This gene-gene interaction may contribute to inter-individual variation in susceptibility to CD. The different findings between these studies may reflect genuine population differences, or it could mean that the specifically tested variants are not directly pathogenic.

CONCLUSION

Epidemiologic data strongly suggest that genetic susceptibility is a major contributing factor to CD. Viewing the association studies, many allelic variants are involved, and great discrepancies can be found. Usually, this is explained by the widespread heterogeneity of CD, and the fact that single genes contribute little to the complex phenotype. Most associations of candidate genes could not be reproduced, didn't aid in understanding pathogenesis, or didn't facilitate diagnosis. Alternatively, genome-wide screening is less biased, and has been proven very useful with the identification of *IBD1* and *CARD15*, which has been replicated many times. Several aspects of future linkage and association analysis should be taken into account: 1) clinical classifications of individuals should be more conform; 2) sample size is a critical determinant for false-negative errors in such studies, especially when stratifying to phenotype or genotype; 3) association of genes is usually performed with selected variants of the gene, instead of sequencing the whole gene. In addition, knowledge of the haplotype structure surrounding genes of interest, allows for common variation in a gene to be tested with more statistical power, even if the causal variant within the shared haplotype block has not been identified (Daly *et al.* 2001).

In 1997, a group of scientists and clinicians founded the IBD International Genetic Consortium (IBDIGC), now a group of twelve research groups involved in the study of genes that are implicated in IBD (Cavanaugh 2003). They collaboratively study large numbers of well-documented families for linkage. Similarly, small-scale pooling of populations may significantly increase the power of linkage and association studies. Finally, the correct choice of candidate genes is another major issue. The IBD loci that have been repeatedly verified are to be selected preferentially, especially if functional data regarding the gene is available through other analysis such as expression data.

REFERENCES

Abbott, D. W., A. Wilkins, J. M. Asara*, et al.* (2004). "The Crohn's disease protein, NOD2, requires RIP2 in order to induce ubiquitinylation of a novel site on NEMO." Curr Biol.

14(24): 2217-27.

Abreu, M. T., K. D. Taylor, Y. C. Lin*, et al.* (2002). "Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease." Gastroenterology **123**(3): 679-88.

Ahmad, T., A. Armuzzi, M. Bunce*, et al.* (2002). "The molecular classification of the clinical manifestations of Crohn's disease." Gastroenterology. **122**(4): 854-66.

Akolkar, P. N., B. Gulwani-Akolkar, D. Heresbach*, et al.* (1997). "Differences in risk of Crohn's disease in offspring of mothers and fathers with inflammatory bowel disease." Am J Gastroenterol. **92**(12): 2241-4.

Akolkar, P. N., B. Gulwani-Akolkar, X. Y. Lin*, et al.* (2001). "The IBD1 locus for susceptibility to Crohn's disease has a greater impact in Ashkenazi Jews with early onset disease." Am J Gastroenterol. **96**(4): 1127-32.

Annese, V., A. Latiano, P. Bovio*, et al.* (1999). "Genetic analysis in Italian families with inflammatory bowel disease supports linkage to the IBD1 locus--a GISC study." Eur J Hum Genet. **7**(5): 567-73.

Arbour, N. C., E. Lorenz, B. C. Schutte*, et al.* (2000). "TLR4 mutations are associated with endotoxin hyporesponsiveness in humans." Nat Genet. **25**(2): 187-91.

Arnott, I. D., E. R. Nimmo, H. E. Drummond*, et al.* (2004). "NOD2/CARD15, TLR4 and CD14 mutations in Scottish and Irish Crohn's disease patients: evidence for genetic heterogeneity within Europe?" Genes Immun. **5**(5): 417-25.

Autschbach, F., S. Eisold, U. Hinz*, et al.* (2005). "High prevalence of Mycobacterium avium subspecies paratuberculosis IS900 DNA in gut tissues from individuals with Crohn's disease." Gut. **54**(7): 944-9.

Bairead, E., D. L. Harmon, A. M. Curtis*, et al.* (2003). "Association of NOD2 with Crohn's disease in a homogenous Irish population." Eur J Hum Genet. **11**(3): 237-44.

Barmada, M. M., S. R. Brant, D. L. Nicolae*, et al.* (2004). "A genome scan in 260 inflammatory bowel disease-affected relative pairs." Inflamm Bowel Dis. **10**(5): 513-20.

Barnich, N., J. E. Aguirre, H. C. Reinecker*, et al.* (2005a). "Membrane recruitment of NOD2 in intestinal epithelial cells is essential for nuclear factor-{kappa}B activation in muramyl dipeptide recognition." J Cell Biol. **170**(1): 21-6.

Barnich, N., T. Hisamatsu, J. E. Aguirre*, et al.* (2005b). "GRIM-19 interacts with nucleotide oligomerization domain 2 and serves as downstream effector of anti-bacterial function in intestinal epithelial cells." J Biol Chem. **280**(19): 19021-6.

Beaugerie, L., N. Massot, F. Carbonnel*, et al.* (2001). "Impact of cessation of smoking on the course of ulcerative colitis." Am J Gastroenterol. **96**(7): 2113-6.

Berrebi, D., R. Maudinas, J. P. Hugot*, et al.* (2003). "Card15 gene overexpression in

mononuclear and epithelial cells of the inflamed Crohn's disease colon." Gut. **52**(6): 840-6.

Bertin, J., W. J. Nir, C. M. Fischer*, et al.* (1999). "Human CARD4 protein is a novel CED-4/Apaf-1 cell death family member that activates NF-kappaB." J Biol Chem. **274**(19): 12955-8.

Bonen, D. K., Y. Ogura, D. L. Nicolae*, et al.* (2003). "Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan." Gastroenterology **124**(1): 140-6.

Braat, H., P. Stokkers, T. Hommes*, et al.* (2005). "Consequence of functional Nod2 and Tlr4 mutations on gene transcription in Crohn's disease patients." J Mol Med. **83**(8): 601-9.

Brant, S. R., Y. Fu, C. T. Fields*, et al.* (1998). "American families with Crohn's disease have strong evidence for linkage to chromosome 16 but not chromosome 12." Gastroenterology. **115**(5): 1056-61.

Burckhardt, G. and N. A. Wolff (2000). "Structure of renal organic anion and cation transporters." Am J Physiol Renal Physiol. **278**(6): F853-66.

Cavanaugh, J. (2001). "International collaboration provides convincing linkage replication in complex disease through analysis of a large pooled data set: Crohn disease and chromosome 16." Am J Hum Genet. **68**(5): 1165-71.

Cavanaugh, J. A. (2003). "IBD International Genetics Consortium: international cooperation making sense of complex disease." Inflamm Bowel Dis. **9**(3): 190-3.

Cavanaugh, J. A., D. F. Callen, S. R. Wilson*, et al.* (1998). "Analysis of Australian Crohn's disease pedigrees refines the localization for susceptibility to inflammatory bowel disease on chromosome 16." Ann Hum Genet. **62**(Pt 4): 291-8.

Chamaillard, M., S. E. Girardin, J. Viala*, et al.* (2003). "Nods, Nalps and Naip: intracellular regulators of bacterial-induced inflammation." Cell Microbiol. **5**(9): 581-92.

Chen, C. M., Y. Gong, M. Zhang*, et al.* (2004). "Reciprocal cross-talk between Nod2 and TAK1 signaling pathways." J Biol Chem. **279**(24): 25876-82.

Cho, J. H., K. B.S. and H. S.B. (1997). "Confirmation of a susceptibility locus for Crohn's disease on chromosome 16." Inflamm Bowel Dis **3**: 186-190.

Cho, J. H., D. L. Nicolae, L. H. Gold*, et al.* (1998). "Identification of novel susceptibility loci for inflammatory bowel disease on chromosomes 1p, 3q, and 4q: evidence for epistasis between 1p and IBD1." Proc Natl Acad Sci U S A. **95**(13): 7502-7.

Colhoun, H. M., P. M. McKeigue and G. Davey Smith (2003). "Problems of reporting genetic associations with complex outcomes." Lancet. **361**(9360): 865-72.

Cosnes, J., L. Beaugerie, F. Carbonnel*, et al.* (2001). "Smoking cessation and the course of Crohn's disease: an intervention study." Gastroenterology. **120**(5): 1093-9.

Croucher, P. J., S. Mascheretti, J. Hampe*, et al.* (2003). "Haplotype structure and association to Crohn's disease of CARD15 mutations in two ethnically divergent populations." Eur J Hum Genet. **11**(1): 6-16.

Curran, M. E., K. F. Lau, J. Hampe*, et al.* (1998). "Genetic analysis of inflammatory bowel disease in a large European cohort supports linkage to chromosomes 12 and 16." Gastroenterology. **115**(5): 1066-71.

Cuthbert, A. P., S. A. Fisher, M. M. Mirza*, et al.* (2002). "The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease." Gastroenterology. **122**(4): 867-74.

Daly, M. J., A. V. Pearce, L. Farwell*, et al.* (2005). "Association of DLG5 R30Q variant with inflammatory bowel disease." Eur J Hum Genet. **13**(7): 835-9.

Daly, M. J., J. D. Rioux, S. F. Schaffner*, et al.* (2001). "High-resolution haplotype structure in the human genome." Nat Genet. **29**(2): 229-32.

Darfeuille-Michaud, A., J. Boudeau, P. Bulois*, et al.* (2004). "High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease." Gastroenterology. **127**(2): 412-21.

Dixon, M. S., C. Golstein, C. M. Thomas*, et al.* (2000). "Genetic complexity of pathogen perception by plants: the example of Rcr3, a tomato gene required specifically by Cf-2." Proc Natl Acad Sci U S A. **97**(16): 8807-14.

Duerr, R. H., M. M. Barmada, L. Zhang*, et al.* (2000). "High-density genome scan in Crohn disease shows confirmed linkage to chromosome 14q11-12." Am J Hum Genet. **66**(6): 1857-62.

Erridge, C., J. Stewart and I. R. Poxton (2003). "Monocytes heterozygous for the Asp299Gly and Thr399Ile mutations in the Toll-like receptor 4 gene show no deficit in lipopolysaccharide signalling." J Exp Med. **197**(12): 1787-91.

Esters, N., M. Pierik, K. van Steen*, et al.* (2004). "Transmission of CARD15 (NOD2) variants within families of patients with inflammatory bowel disease." Am J Gastroenterol. **99**(2): 299-305.

Farrokhyar, F., E. T. Swarbrick and E. J. Irvine (2001). "A critical review of epidemiological studies in inflammatory bowel disease." Scand J Gastroenterol. **36**(1): 2-15.

Fidder, H. H., B. Avidan, M. Lahav*, et al.* (2003). "Clinical and demographic characterization of Jewish Crohn's disease patients in Israel." J Clin Gastroenterol **36**(1): 8-12.

Franchimont, D., S. Vermeire, H. El Housni*, et al.* (2004). "Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis." Gut. **53**(7): 987-92.

Gazouli, M., G. Mantzaris, A. Kotsinas*, et al.* (2005). "Association between polymorphisms in the Toll-like receptor 4, CD14, and CARD15/NOD2 and inflammatory bowel disease in

the Greek population." World J Gastroenterol. **11**(5): 681-5.

Giallourakis, C., M. Stoll, K. Miller*, et al.* (2003). "IBD5 is a general risk factor for inflammatory bowel disease: replication of association with Crohn disease and identification of a novel association with ulcerative colitis." Am J Hum Genet. **73**(1): 205-11.

Gibson, P. R. and S. J. Shepherd (2005). "Personal view: food for thought--western lifestyle and susceptibility to Crohn's disease. The FODMAP hypothesis." Aliment Pharmacol Ther. **21**(12): 1399-409.

Girardin, S. E., I. G. Boneca, J. Viala*, et al.* (2003). "Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection." J Biol Chem. **278**(11): 8869-72.

Goodgame, R. W., K. Kimball, S. Akram*, et al.* (2001). "Randomized controlled trial of clarithromycin and ethambutol in the treatment of Crohn's disease." Aliment Pharmacol Ther. **15**(12): 1861-6.

Griga, T., W. Klein, J. T. Epplen*, et al.* (2005). "CD14 expression on monocytes and soluble CD14 plasma levels in correlation to the promotor polymorphism of the endotoxin receptor CD14 gene in patients with inactive Crohn's disease." Hepatogastroenterology. **52**(63): 808-11.

Gutierrez, O., C. Pipaon, N. Inohara*, et al.* (2002). "Induction of Nod2 in myelomonocytic and intestinal epithelial cells via nuclear factor-kappa B activation." J Biol Chem. **277**(44): 41701-5.

Halfvarson, J., L. Bodin, C. Tysk*, et al.* (2003). "Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics." Gastroenterology **124**(7): 1767-73.

Halme, L., U. Turunen, P. Paavola-Sakki*, et al.* (2004). "CARD15 frameshift mutation in patients with Crohn's disease is associated with immune dysregulation." Scand J Gastroenterol. **39**(12): 1243-9.

Hampe, J., J. Grebe, S. Nikolaus*, et al.* (2002). "Association of NOD2 (CARD 15) genotype with clinical course of Crohn's disease: a cohort study." Lancet. **359**(9318): 1661-5.

Hampe, J., S. Schreiber, S. H. Shaw*, et al.* (1999). "A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort." Am J Hum Genet. **64**(3): 808-16.

Helio, T., L. Halme, M. Lappalainen*, et al.* (2003). "CARD15/NOD2 gene variants are associated with familially occurring and complicated forms of Crohn's disease." Gut. **52**(4): 558-62.

Hisamatsu, T., M. Suzuki, H. C. Reinecker*, et al.* (2003). "CARD15/NOD2 functions as an antibacterial factor in human intestinal epithelial cells." Gastroenterology. **124**(4): 993-1000.

Ho, P., I. N. Bruce, A. Silman*, et al.* (2005). "Evidence for common genetic control in pathways of inflammation for Crohn's disease and psoriatic arthritis." Arthritis Rheum. **52**(11): 3596-602.

Hugot, J. P., C. Alberti, D. Berrebi*, et al.* (2003). "Crohn's disease: the cold chain hypothesis." Lancet. **362**(9400): 2012-5.

Hugot, J. P., M. Chamaillard, H. Zouali*, et al.* (2001). "Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease." Nature **411**(6837): 599-603.

Hugot, J. P., P. Laurent-Puig, C. Gower-Rousseau*, et al.* (1996). "Mapping of a susceptibility locus for Crohn's disease on chromosome 16." Nature **379**(6568): 821-3.

Humbert, P., S. Russell and H. Richardson (2003). "Dlg, Scribble and Lgl in cell polarity, cell proliferation and cancer." Bioessays. **25**(6): 542-53.

Inohara, Chamaillard, C. McDonald*, et al.* (2005). "NOD-LRR proteins: role in host-microbial interactions and inflammatory disease." Annu Rev Biochem. **74**: 355-83.

Inohara, N., T. Koseki, L. del Peso*, et al.* (1999). "Nod1, an Apaf-1-like activator of caspase-9 and nuclear factor-kappaB." J Biol Chem. **274**(21): 14560-7.

Inohara, N. and G. Nunez (2003). "NODs: intracellular proteins involved in inflammation and apoptosis." Nat Rev Immunol. **3**(5): 371-82.

Inohara, N., Y. Ogura, A. Fontalba*, et al.* (2003). "Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease." J Biol Chem. **278**(8): 5509-12.

Inoue, N., K. Tamura, Y. Kinouchi*, et al.* (2002). "Lack of common NOD2 variants in Japanese patients with Crohn's disease." Gastroenterology. **123**(1): 86-91.

Jump, R. L. and A. D. Levine (2004). "Mechanisms of natural tolerance in the intestine: implications for inflammatory bowel disease." Inflamm Bowel Dis. **10**(4): 462-78.

Klein, W., A. Tromm, T. Griga*, et al.* (2003). "Interaction of polymorphisms in the CARD15 and CD14 genes in patients with Crohn disease." Scand J Gastroenterol. **38**(8): 834-6.

Klein, W., A. Tromm, T. Griga*, et al.* (2002). "A polymorphism in the CD14 gene is associated with Crohn disease." Scand J Gastroenterol. **37**(2): 189-91.

Kobayashi, K., N. Inohara, L. D. Hernandez*, et al.* (2002). "RICK/Rip2/CARDIAK mediates signalling for receptors of the innate and adaptive immune systems." Nature. **416**(6877): 194-9.

Kobayashi, K. S., M. Chamaillard, Y. Ogura*, et al.* (2005). "Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract." Science. **307**(5710): 731-4.

Kramer, M., M. G. Netea, D. J. de Jong*, et al.* (2006). "Impaired dendritic cell function in Crohn's disease patients with NOD2 3020insC mutation." J Leukoc Biol **3**: 3.

Lakatos, P. L., L. Lakatos, F. Szalay, *et al.* (2005). "Toll-like receptor 4 and NOD2/CARD15 mutations in Hungarian patients with Crohn's disease: phenotype-genotype correlations." World J Gastroenterol. **11**(10): 1489-95.

Lala, S., Y. Ogura, C. Osborne, *et al.* (2003). "Crohn's disease and the NOD2 gene: a role for paneth cells." Gastroenterology. **125**(1): 47-57.

Leong, R. W., A. Armuzzi, T. Ahmad, *et al.* (2003). "NOD2/CARD15 gene polymorphisms and Crohn's disease in the Chinese population." Aliment Pharmacol Ther. **17**(12): 1465-70.

Lesage, S., H. Zouali, J. P. Cezard, *et al.* (2002). "CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease." Am J Hum Genet. **70**(4): 845-57.

Leung, E., J. Hong, A. G. Fraser, *et al.* (2005). "Polymorphisms of CARD15/NOD2 and CD14 genes in New Zealand Crohn's disease patients." Immunol Cell Biol. **83**(5): 498-503.

Levine, A., A. Karban, R. Eliakim, *et al.* (2005). "A polymorphism in the TNF-alpha promoter gene is associated with pediatric onset and colonic location of Crohn's disease." Am J Gastroenterol. **100**(2): 407-13.

Li, J., T. Moran, E. Swanson, *et al.* (2004). "Regulation of IL-8 and IL-1beta expression in Crohn's disease associated NOD2/CARD15 mutations." Hum Mol Genet. **13**(16): 1715-25.

Linderson, Y., F. Bresso, E. Buentke, *et al.* (2005). "Functional interaction of CARD15/NOD2 and Crohn's disease-associated TNFalpha polymorphisms." Int J Colorectal Dis. **20**(4): 305-11.

Logan, R. F. (1999). "Smoking, use of oral contraceptives, and medical induction of remission were risk factors for relapse in Crohn's disease." Gut **44**(3): 311-2.

Ma, Y., J. D. Ohmen, Z. Li, *et al.* (1999). "A genome-wide search identifies potential new susceptibility loci for Crohn's disease." Inflamm Bowel Dis. **5**(4): 271-8.

Maeda, S., L. C. Hsu, H. Liu, *et al.* (2005). "Nod2 mutation in Crohn's disease potentiates NF-kappaB activity and IL-1beta processing." Science. **307**(5710): 734-8.

Miceli-Richard, C., S. Lesage, M. Rybojad, *et al.* (2001). "CARD15 mutations in Blau syndrome." Nat Genet **29**(1): 19-20.

Monsen, U., O. Bernell, C. Johansson, *et al.* (1991). "Prevalence of inflammatory bowel disease among relatives of patients with Crohn's disease." Scand J Gastroenterol **26**(3): 302-6.

Nakamura, H., T. Sudo, H. Tsuiki, *et al.* (1998). "Identification of a novel human homolog of the Drosophila dlg, P-dlg, specifically expressed in the gland tissues and interacting with p55." FEBS Lett. **433**(1-2): 63-7.

Negoro, K., Y. Kinouchi, N. Hiwatashi, *et al.* (1999). "Crohn's disease is associated with novel polymorphisms in the 5'-flanking region of the tumor necrosis factor gene."

Gastroenterology. **117**(5): 1062-8.

Negoro, K., D. P. McGovern, Y. Kinouchi*, et al.* (2003). "Analysis of the IBD5 locus and potential gene-gene interactions in Crohn's disease." Gut. **52**(4): 541-6.

Netea, M. G., G. Ferwerda, D. J. de Jong*, et al.* (2005). "NOD2 3020insC mutation and the pathogenesis of Crohn's disease: impaired IL-1beta production points to a loss-of-function phenotype." Neth J Med. **63**(8): 305-8.

Netea, M. G., B. J. Kullberg, D. J. de Jong*, et al.* (2004). "NOD2 mediates anti-inflammatory signals induced by TLR2 ligands: implications for Crohn's disease." Eur J Immunol. **34**(7): 2052-9.

Newman, W. G., X. Gu, R. F. Wintle*, et al.* (2006). "DLG5 variants contribute to Crohn disease risk in a Canadian population." Hum Mutat **31**: 31.

Noble, C. L., E. R. Nimmo, H. Drummond*, et al.* (2005a). "The contribution of OCTN1/2 variants within the IBD5 locus to disease susceptibility and severity in Crohn's disease." Gastroenterology. **129**(6): 1854-64.

Noble, C. L., E. R. Nimmo, H. Drummond*, et al.* (2005b). "DLG5 variants do not influence susceptibility to inflammatory bowel disease in the Scottish population." Gut. **54**(10): 1416-20.

Odes, H. S., D. Fraser and J. Krawiec (1989). "Inflammatory bowel disease in migrant and native Jewish populations of southern Israel." Scand J Gastroenterol Suppl **170**: 36-8; discussion 50-5.

Ogura, Y., D. K. Bonen, N. Inohara*, et al.* (2001a). "A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease." Nature **411**(6837): 603-6.

Ogura, Y., N. Inohara, A. Benito*, et al.* (2001b). "Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB." J Biol Chem. **276**(7): 4812-8.

Ogura, Y., S. Lala, W. Xin*, et al.* (2003). "Expression of NOD2 in Paneth cells: a possible link to Crohn's ileitis." Gut. **52**(11): 1591-7.

Oh, H. M., H. J. Lee, G. S. Seo*, et al.* (2005). "Induction and localization of NOD2 protein in human endothelial cells." Cell Immunol. **237**(1): 37-44.

Ohmen, J. D., H. Y. Yang, K. K. Yamamoto*, et al.* (1996). "Susceptibility locus for inflammatory bowel disease on chromosome 16 has a role in Crohn's disease, but not in ulcerative colitis." Hum Mol Genet. **5**(10): 1679-83.

Ouellette, A. J. and C. L. Bevins (2001). "Paneth cell defensins and innate immunity of the small bowel." Inflamm Bowel Dis. **7**(1): 43-50.

Paavola-Sakki, P., V. Ollikainen, T. Helio*, et al.* (2003). "Genome-wide search in Finnish families with inflammatory bowel disease provides evidence for novel susceptibility loci." Eur J Hum Genet. **11**(2): 112-20.

Peeters, H., S. Bogaert, D. Laukens*, et al.* (2006). "*CARD15* variants determine a disturbed response of monocytes to adherent-invasive *Escherichia coli LF82* in Crohn's disease." *Submitted to IBD*.

Peeters, M., H. Nevens, F. Baert*, et al.* (1996). "Familial aggregation in Crohn's disease: increased age-adjusted risk and concordance in clinical characteristics." Gastroenterology. **111**(3): 597-603.

Peltekova, V. D., R. F. Wintle, L. A. Rubin*, et al.* (2004). "Functional variants of OCTN cation transporter genes are associated with Crohn disease." Nat Genet. **36**(5): 471-5.

Peters, K. E., N. J. O'Callaghan and J. A. Cavanaugh (2005). "Lack of association of the CD14 promoter polymorphism--159C/T with Caucasian inflammatory bowel disease." Scand J Gastroenterol. **40**(2): 194-7.

Podolsky, D. K. (2002). "Inflammatory bowel disease." N Engl J Med. **347**(6): 417-29.

Present, D. H., P. Rutgeerts, S. Targan*, et al.* (1999). "Infliximab for the treatment of fistulas in patients with Crohn's disease." N Engl J Med. **340**(18): 1398-405.

Price, A. B. (1978). "Overlap in the spectrum of non-specific inflammatory bowel disease--'colitis indeterminate'." J Clin Pathol. **31**(6): 567-77.

Rahman, P., S. Bartlett, F. Siannis*, et al.* (2003). "CARD15: a pleiotropic autoimmune gene that confers susceptibility to psoriatic arthritis." Am J Hum Genet. **73**(3): 677-81.

Rioux, J. D., M. J. Daly, M. S. Silverberg*, et al.* (2001). "Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease." Nat Genet. **29**(2): 223-8.

Rioux, J. D., M. S. Silverberg, M. J. Daly*, et al.* (2000). "Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci." Am J Hum Genet. **66**(6): 1863-70.

Rogler, G., T. Andus, E. Aschenbrenner*, et al.* (1997). "Alterations of the phenotype of colonic macrophages in inflammatory bowel disease." Eur J Gastroenterol Hepatol. **9**(9): 893-9.

Rosenstiel, P., M. Fantini, K. Brautigam*, et al.* (2003). "TNF-alpha and IFN-gamma regulate the expression of the NOD2 (CARD15) gene in human intestinal epithelial cells." Gastroenterology. **124**(4): 1001-9.

Rutgeerts, P., K. Goboes, M. Peeters*, et al.* (1991). "Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum." Lancet. **338**(8770): 771-4.

Ryan, P., R. G. Kelly, G. Lee*, et al.* (2004). "Bacterial DNA within granulomas of patients with Crohn's disease--detection by laser capture microdissection and PCR." Am J Gastroenterol **99**(8): 1539-43.

Sartor, R. B. (1994). "Cytokines in intestinal inflammation: pathophysiological and clinical considerations." Gastroenterology. **106**(2): 533-9.

Sartor, R. B. (2003). "Targeting enteric bacteria in treatment of inflammatory bowel diseases: why, how, and when." Curr Opin Gastroenterol. **19**(4): 358-65.

Satsangi, J., C. Grootscholten, H. Holt*, et al.* (1996a). "Clinical patterns of familial inflammatory bowel disease." Gut **38**(5): 738-41.

Satsangi, J., M. Parkes, E. Louis*, et al.* (1996b). "Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12." Nat Genet. **14**(2): 199-202.

Saxelin, M., S. Tynkkynen, T. Mattila-Sandholm*, et al.* (2005). "Probiotic and other functional microbes: from markets to mechanisms." Curr Opin Biotechnol **16**(2): 204-11.

Schreiber, S., S. Nikolaus and J. Hampe (1998). "Activation of nuclear factor kappa B inflammatory bowel disease." Gut. **42**(4): 477-84.

Schwede, T., J. Kopp, N. Guex*, et al.* (2003). "SWISS-MODEL: An automated protein homology-modeling server." Nucleic Acids Res. **31**(13): 3381-5.

Shaw, S. H., J. Hampe, R. White*, et al.* (2003). "Stratification by CARD15 variant genotype in a genome-wide search for inflammatory bowel disease susceptibility loci." Hum Genet. **113**(6): 514-21.

Shi, H. N. and A. Walker (2004). "Bacterial colonization and the development of intestinal defences." Can J Gastroenterol. **18**(8): 493-500.

Stoll, M., B. Corneliussen, C. M. Costello*, et al.* (2004). "Genetic variation in DLG5 is associated with inflammatory bowel disease." Nat Genet. **36**(5): 476-80.

Swidsinski, A., A. Ladhoff, A. Pernthaler*, et al.* (2002). "Mucosal flora in inflammatory bowel disease." Gastroenterology. **122**(1): 44-54.

Tada, H., S. Aiba, K. Shibata*, et al.* (2005). "Synergistic effect of Nod1 and Nod2 agonists with toll-like receptor agonists on human dendritic cells to generate interleukin-12 and T helper type 1 cells." Infect Immun. **73**(12): 7967-76.

Tanabe, T., M. Chamaillard, Y. Ogura*, et al.* (2004). "Regulatory regions and critical residues of NOD2 involved in muramyl dipeptide recognition." Embo J. **23**(7): 1587-97.

Targan, S. R., S. B. Hanauer, S. J. van Deventer*, et al.* (1997). "A short-term study of chimeric monoclonal antibody cA2 to tumor necrosis factor alpha for Crohn's disease. Crohn's Disease cA2 Study Group." N Engl J Med. **337**(15): 1029-35.

Tenesa, A., C. Noble, J. Satsangi*, et al.* (2006). "Association of DLG5 and inflammatory bowel disease across populations." Eur J Hum Genet **4**: 4.

Thompson, N. P., R. Driscoll, R. E. Pounder*, et al.* (1996). "Genetics versus environment in inflammatory bowel disease: results of a British twin study." Bmj **312**(7023): 95-6.

Timmer, A. (2003). "Environmental influences on inflammatory bowel disease manifestations.

Lessons from epidemiology." Dig Dis **21**(2): 91-104.

Torok, H. P., J. Glas, L. Tonenchi*, et al.* (2005). "Polymorphisms in the DLG5 and OCTN cation transporter genes in Crohn's disease." Gut. **54**(10): 1421-7.

Torok, H. P., J. Glas, L. Tonenchi*, et al.* (2004). "Polymorphisms of the lipopolysaccharide-signaling complex in inflammatory bowel disease: association of a mutation in the Toll-like receptor 4 gene with ulcerative colitis." Clin Immunol. **112**(1): 85-91.

Tysk, C., E. Lindberg, G. Jarnerot*, et al.* (1988). "Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking." Gut **29**(7): 990-6.

Van Gossum, A., M. Adler, M. De Reuck*, et al.* (1996). "Epidemiology of inflammatory bowel disease in Brussels' area (1992-1993)." Acta Gastroenterol Belg. **59**(1): 7-9.

van Heel, D. A., B. M. Dechairo, G. Dawson*, et al.* (2003). "The IBD6 Crohn's disease locus demonstrates complex interactions with CARD15 and IBD5 disease-associated variants." Hum Mol Genet. **12**(20): 2569-75..

van Heel, D. A., S. Ghosh, M. Butler*, et al.* (2005a). "Muramyl dipeptide and toll-like receptor sensitivity in NOD2-associated Crohn's disease." Lancet. **365**(9473): 1794-6.

van Heel, D. A., S. Ghosh, K. A. Hunt*, et al.* (2005b). "Synergy between TLR9 and NOD2 innate immune responses is lost in genetic Crohn's disease." Gut. **54**(11): 1553-7.

van Heel, D. A., I. A. Udalova, A. P. De Silva*, et al.* (2002). "Inflammatory bowel disease is associated with a TNF polymorphism that affects an interaction between the OCT1 and NF(-kappa)B transcription factors." Hum Mol Genet. **11**(11): 1281-9.

Van Kruiningen, H. J., M. Joossens, S. Vermeire*, et al.* (2005). "Environmental factors in familial Crohn's disease in Belgium." Inflamm Bowel Dis **11**(4): 360-5.

Vermeire, S., M. Peeters and P. Rutgeerts (2000). "Diagnostic approach to IBD." Hepatogastroenterology. **47**(31): 44-8.

Vermeire, S., P. Rutgeerts, K. Van Steen*, et al.* (2004). "Genome wide scan in a Flemish inflammatory bowel disease population: support for the IBD4 locus, population heterogeneity, and epistasis." Gut **53**(7): 980-6.

Vermeire, S., J. Satsangi, M. Peeters*, et al.* (2001). "Evidence for inflammatory bowel disease of a susceptibility locus on the X chromosome." Gastroenterology. **120**(4): 834-40.

Vermeire, S., G. Wild, K. Kocher*, et al.* (2002). "CARD15 Genetic Variation in a Quebec Population: Prevalence, Genotype- Phenotype Relationship, and Haplotype Structure." Am J Hum Genet **71**(1): 74-83.

von Aulock, S., N. W. Schroder, K. Gueinzius*, et al.* (2003). "Heterozygous toll-like receptor 4 polymorphism does not influence lipopolysaccharide-induced cytokine release in human whole blood." J Infect Dis. **188**(6): 938-43.

Voss, E., J. Wehkamp, K. Wehkamp*, et al.* (2005). "NOD2/CARD15 mediates induction of the antimicrobial peptide human beta-defensin-2." J Biol Chem **30**: 30.

Wakabayashi, M., T. Ito, M. Mitsushima*, et al.* (2003). "Interaction of lp-dlg/KIAA0583, a membrane-associated guanylate kinase family protein, with vinexin and beta-catenin at sites of cell-cell contact." J Biol Chem. **278**(24): 21709-14.

Waller, S., M. Tremelling, F. Bredin*, et al.* (2005). "Evidence for association of OCTN genes and IBD5 with ulcerative colitis." Gut **16**: 16.

Watanabe, T., A. Kitani, P. J. Murray*, et al.* (2004). "NOD2 is a negative regulator of Toll-like receptor 2-mediated T helper type 1 responses." Nat Immunol. **5**(8): 800-8..

Wehkamp, J., J. Harder, M. Weichenthal*, et al.* (2004). "NOD2 (CARD15) mutations in Crohn's disease are associated with diminished mucosal alpha-defensin expression." Gut. **53**(11): 1658-64.

Wells, R. W. and M. G. Blennerhassett (2005). "The increasing prevalence of Crohn's disease in industrialized societies: the price of progress?" Can J Gastroenterol. **19**(2): 89-95.

Westendorp, R. G., J. A. Langermans, T. W. Huizinga*, et al.* (1997). "Genetic influence on cytokine production and fatal meningococcal disease." Lancet. **349**(9046): 170-3.

Wright, S. D., R. A. Ramos, P. S. Tobias*, et al.* (1990). "CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein." Science. **249**(4975): 1431-3.

Chapter 2

A FUNCTIONAL PROMOTER POLYMORPHISM IN *CARD15*

Debby Laukens[1], Michaela Loos[1], Jurgen Del-Favero[2], Sara Bogaert[3], Harald Peeters[3], Martine De Vos[3] and Pieter Rottiers[1]

[1]Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium
[2]Department of Molecular Genetics, University of Antwerp, VIB, Antwerpen, Belgium
[3]Department of Gastroenterology, Ghent University Hospital, Gent, Belgium

ABSTRACT

Background & aims: Three polymorphisms within the *CARD15* gene have been independently associated with Crohn's disease (CD). Although this association is strong, an average of 60 to 70% of CD patients do not have any of these three mutations. Furthermore, they do not account for all of the linkage observed between CD and the *IBD1* locus. Since CARD15 is highly up-regulated during inflammation, we looked for polymorphisms in the *CARD15* promoter region.
Methods: The promoter region of *CARD15* was sequenced in 28 Flemish CD patients, and a polymorphism was genotyped in 104 CD patients and 60 controls. An *in vitro* gene reporter assay was used to assess the functionality of the promoter polymorphism in response to tumor necrosis factor alpha (TNF). Public HapMap data from CEPH trios were used to determine the haplotype structure surrounding *CARD15*.
Results: Sequencing of the promoter region revealed a c.-59G>A polymorphism in the 5' untranslated region (5'UTR), contained within a conserved potential binding site for the E2F transcription factor. This polymorphism was significantly associated with CD. However, in 99% of the cases, the CD-associated *CARD15* mutations occur in the c-59G>A background, suggesting that they are linked. In addition, the infrequent c.-59A allele was less responsive to TNF stimulation. The complete *CARD15* gene is located within a single haplotype block.
Conclusions: A functional promoter polymorphism, linked to the common CD-associated mutations, was identified in the *CARD15* 5'UTR. The complete *CARD15* gene is located in a

single haplotype block, and we are currently determining the haplotype structure in our patient population to evaluate the transmission of the *CARD15* promoter polymorphism.

INTRODUCTION

Crohn's disease (CD, MIM 266600) is an inflammatory condition of the gastrointestinal tract. The disease is believed to result from an interaction between environmental factors and genetic predisposition. Since 1996, a locus at 16q12 has been repeatedly associated with increased susceptibility to CD (Hugot *et al.* 1996). In 2001, two groups independently identified the first strongly CD-associated gene called *NOD2*, later referred to as *CARD15* (Hugot *et al.* 2001; Ogura *et al.* 2001a). CARD15 is a cytosolic receptor for bacterial components and initiates an inflammatory response following bacterial challenge. Both groups found the same gene using a different approach. Hugot and co-workers systematically refined the susceptibility locus of chromosome 16. They characterized *CARD15* within this region, and sequenced the 11 exons of this gene in 50 unrelated CD patients. Three haplotypes with preferential transmission to affected individuals were identified, each containing one rare allele of single nucleotide polymorphism (SNP) 8 (c.2104C>T), SNP12 (c.2722G>C) or SNP13 (c.3020insC) in the context of a common background. SNP8 and SNP12 result in non-conservative amino acid substitutions R702W and G908R respectively. SNP13 is a frameshift mutation (1007fs), which truncates the tenth leucine-rich repeat of CARD15. On the other hand, Ogura and colleagues used the candidate gene approach, which is based on an intelligent search for genes within regions of linkage. They identified the SNP13 mutation, which was preferentially transmitted, and assessed this SNP in a large case-control study. The association of the three mutations have since been confirmed in many studies. In Europe, they are found in approximately one third of CD patients, especially in those with ileal disease (Cuthbert *et al.* 2002).

The observed linkage of CD to chromosome 16 can not be entirely explained by the present associations (Hugot *et al.* 2001). Furthermore, linkage to chromosome 16 was still observed in CD patients not possessing one of the three *CARD15* mutations (Hampe *et al.* 2002; Shaw *et al.* 2003; van Heel *et al.* 2003). Thus, other variants of this gene or variations in neighbouring genes on chromosome 16 may be involved in CD susceptibility.

The expression of the CARD15 protein in normal colon is restricted to scattered mononuclear cells in the lamina propria (Berrebi *et al.* 2003), and in paneth cells in the ileum (Ogura *et al.* 2003). In the CD colon, however, the number of positive cells is correlated with the inflammatory infiltrate. In severely inflamed CD colon samples, intestinal epithelial cells were

also CARD15 positive. Together with the observation that CARD15 is up-regulated in acute appendicitis, this favours the hypothesis that CARD15 is induced upon inflammatory stimuli. Furthermore, it was shown that tumor necrosis factor alpha (TNF), a potent pro-inflammatory cytokine, up-regulates CARD15 via NFκB in epithelial cell lines (Rosenstiel *et al.* 2003). In accordance, two functional NFκB binding sites were found within the *CARD15* promoter. Therefore, CARD15 expression regulation could be an important mechanism in innate immune response in intestinal epithelial cells. We wondered whether polymorphisms are present within the promoter of *CARD15*, which might influence their expression. We sequenced a part of the promoter region in *CARD15*, and found a promoter polymorphism in the 5' untranslated region (5'UTR), c.-59G>A, located within a DNA stretch that potentially binds the E2F transcription factor. The functionality of both alleles was tested in an *in vitro* gene reporter system.

MATERIALS AND METHODS

Sequencing and genotyping of c.-59G>A polymorphism

Genomic DNA was extracted from whole blood using the Qiagen blood and cell culture DNA kit. The promoter region of *CARD15* was PCR-amplified in 28 Flemish CD patients using forward primer: 5'-GGC CTG TCC CCT CGT GAA TG-3' and reverse primer: 5'-GTC GCG GCC AAG GAT GAA AG-3'. This product was sequenced with the BigDye Terminator v3.1 Cycle Sequencing kit and analyzed on an ABI3700 analyzer (Applied Biosystems).

The c.-59G>A polymorphism was genotyped by RFLP-PCR in 104 unrelated CD patients and 60 healthy controls. A fragment containing the polymorphism was amplified with the following primers: forward: 5'-GGC GGA GGT TGG AGT TGA AAA TAA-3'; reverse: 5'-GGA AGC CAG GAT CTA AGG TA-3'. The PCR product was precipitated with SeeDNA (Amersham Biosciences), digested with *Bsm*I, and restriction fragments were analyzed on a 2% agarose gel. The presence of the c.-59A allele creates a *Bsm*I site, resulting in two bands of 1108 and 510 bp, instead of a single 1618 bp band in the presence of the c.-59G allele.

Genotyping for the three CD-associated *CARD15* polymorphisms c.2104C>T, c.2722G>C and c.3020insC was performed using RFLP-PCR as previously described (Laukens *et al.* 2005).

Constructs

Genomic PCR fragments of 155 bp of the promoter region of *CARD15*, starting at base -1 to -155 upstream from the initiation codon, were cloned between the *Sac*I and *Kpn*I sites of the PGL3basic luciferase reporter vector (Promega Corporation Benelux). This generated PGL3-

59A and PGL3-59G constructs. Both constructs were sequence verified. The β-galactosidase expressing pUT651 vector (Eurogentec, Belgium) was used for normalizing transfection efficiencies.

Transfection and gene reporter assay

HEK293T cells (ATCC CRL1573) were seeded in 24-well plates at $5x10^5$ cells per well the day before transfection. Cells were cotransfected with 100 ng pUT651 and 100 ng of PGL3-59G or PGL3-59A in triplicate using the calcium phosphate method. Twenty-four hours posttransfection, cells were incubated with 1000 U/ml recombinant human TNF purified from *E. coli* (produced in-house). After 24 hours, cells were lysed in 150 µl lysis buffer (25 mM Tris pH 7.8, 2 mM DTT, 2mM CDTA, 10% glycerol and 1% Triton X-100) for 15 minutes. Lysates were analyzed for luciferase activity by mixing 50 µl of lysate with 50 µl of luciferase assay buffer (20 mM tricine, 1.07 mM $(MgCO_3)_4M(OH)_2$, 2.67 mM $MgSO_4$, 33.3 mM DTT, 0.1 mM EDTA, 270 µM CoA, 530 µM ATP and 470 µM luciferin). Luciferase activity was measured in a TopCount luminescence counter (Packard). β-galactosidase activity was analyzed by mixing 20 µl of lysate with 160 µl of substrate buffer (60 mM Na_2HPO_4 pH 7.0, 10mM KCl, 1mM β-mercaptoethanol) and 20 µl of chlorophenolred-β-D-galactopyranoside (CPRG) substrate (Roche Diagnostics NV, Belgium). After 30 minutes, β-galactosidase activity was measured at 540 nm in a microplate reader (BioRad Laboratories, Belgium).

Haplotype structure

Public data of the International HapMap Project (http://www.hapmap.org) were used to identify the haplotype structure surrounding *CARD15*, using the CEPH (Centre d'Etude du Polymorphisme Humain) population data, who are all Utah (USA) residents with ancestry from northern and western Europe. The criteria used to assign membership in the CEPH population have not been specified, except that all donors were residents of Utah.

Statistics

Luciferase induction was represented as mean of triplicates + SD. The Student's t test was used to compare groups. Allele frequencies between controls and CD patients were compared using Pearson's χ^2 test. Two-tailed probabilities were calculated, and *P*-values of less than 0.05 were considered statistically significant.

RESULTS

A polymorphism in the promoter region of *CARD15* is linked to CD-associated mutations

To search for polymorphisms in the promoter of *CARD15*, a 425 bp region surrounding the most functional NFκB binding site (Rosenstiel *et al.* 2003) was sequenced in 28 unrelated CD patients. One polymorphism was found, c.-59G>A, and subsequently typed in 104 CD patients and 60 controls (Table 1). An association with CD patients was found for SNP8, SNP12, SNP13 and c.-59G>A, but the latter significance is because it is in linkage disequilibrium (LD) with SNP8 (R^2=0.224, LOD 11.76). If we omit the patients carrying a *CARD15* mutation, the association of c.-59A with CD is lost. Furthermore, no correlation was found with a specific phenotype of CD, e.g. disease location, age of onset or structuring disease (data not shown).

Table 1 Frequencies for SNP8, SNP12, SNP13 and c.-59G>A in our study population

	% rare allele frequency in		
	CD patients (*N*=208)	controls (*N*=120)	*P*-value
SNP8	17.7	4.2	0.002
SNP12	3.8	0.08	0.043
SNP13	10.1	0.03	0.001
c.-59G>A	47.1	26.6	0.003
c.-59G>A no *CARD15*	53.2	42.3	0.119

The c.-59G>A polymorphism is located within a conserved binding site for E2F

A search for transcription factor binding sites using the TRANSFAC database revealed that the c.-59G>A is located within a potential binding site for the E2F transcription factor (Figure 1). To determine whether this binding site is conserved, a search of E2F binding sites in the murine *CARD15* gene and in other members of the Nod1/Apaf1 family (Ogura *et al.* 2001b) was performed. The human *CARD15*, *NOD1* and the murine *CARD15* contain 1 potential binding site for E2F within a 800 bp promoter region upstream of the coding sequence. The *Apaf1* promoter contains 2 of these sites. The conservation of the E2F binding region suggests that regulation of CARD15 by E2F might be an essential mechanism in the function of CARD15.

Figure 1 Part of the promoter region of *CARD15*, containing a functional NFκB site and a potential binding site for E2F. The c.-59G>A polymorphism in the 5'UTR (grey box) is located within the potential binding site for E2F. The E2F consensus sequence is T/C T T C/G G/C C G. The coding sequence of exon 1 is boxed.

The c.-59A allele impairs CARD15 induction by tumor necrosis factor alpha

Because the c.-59G>A polymorphism is located within a potential binding site for the E2F transcription factor, we assessed whether the distinct alleles influence the promoter induction after stimulation with tumor necrosis factor alpha (TNF). Therefore, we generated two plasmids containing the c.-59G (PGL3-59G) or c.-59A allele (PGL3-59A), together with the functional NFκB site in a gene reporter system. After 24 hours of TNF stimulation, luciferase activity induction in HEK293T cells transfected with the PGL3-59A construct was significantly less as compared with PGL3-59G construct (Figure 2). A similar result was obtained when HT29 colonic epithelial cells were used as an acceptor cell (data not shown).

***CARD15* is contained within a single haplotype block**

The haplotype structure surrounding *CARD15* was determined using public CEPH data from the HapMap project. The complete *CARD15* gene is located within a single haplotype block (Figure 3).

CONCLUSION

We identified a polymorphism within the *CARD15* 5'UTR that differs in responsiveness to TNF induction. This polymorphism, currently present in the SNP database as rs.5743266, is in LD with SNP8, a common CD-associated polymorphism.

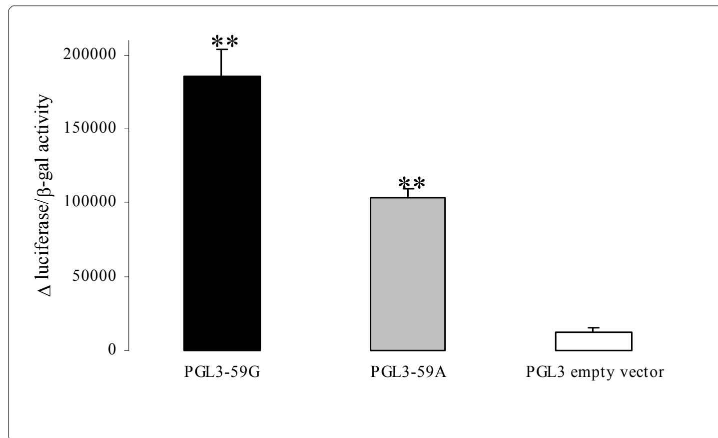


Figure 2 Luciferase induction controlled by a part of the *CARD15* promoter in response to 24 hours of TNF induction is lower in the presence of the c.-59A allele (PGL3-59A) as compared to the c.-59G allele (PGL3-59G).
$^{**}P<0.005$

Because the function of the three polymorphisms in *CARD15* is not completely understood, and because other genetic variants in or close to *CARD15* are supposed to be present, we focused on a promoter polymorphism, which influences the expression of CARD15. Several reports have shown that CARD15 expression is induced by TNF and under inflammatory conditions (Berrebi *et al.* 2003; Rosenstiel *et al.* 2003). It is however not known whether the three CD-associated mutations influence CARD15 expression. In a small cohort, there appeared to be no correlation between *CARD15* mutations and CARD15 expression in the colon. However, only 4 out of 8 CD patients carried one of the *CARD15* mutations (Berrebi *et al.* 2003). Therefore, a detailed expression study of colonic and ileal CARD15 in relation to the common mutations needs to be performed. Moreover, the importance of CARD15 expression has not been extensively studied. It was shown that overexpression of CARD15 in intestinal epithelial cells leads to a higher IL8 secretion in response to lipopolysaccharide (Rosenstiel *et al.* 2003). The IL8 cytokine is a chemotactic protein, which recruits neutrophils, and is a major mediator of the inflammatory response**.** Up-regulation of CARD15 in intestinal inflammation may therefore represent a generic response to bacterial invasion.

The induction of CARD15 transcription in response to inflammation is probably mediated through two functional NFκB sites in the promoter (Rosenstiel *et al.* 2003). The c.-59G>A polymorphism is located within a potential binding site for the E2F transcription factor. E2F is involved in cell cycle progression (Fang and Han 2006). To establish a true functional consequence of the polymorphism, DNA binding of E2F to both alleles needs to be determined. Because the c.-59A allele is the common background of SNP8, SNP12 and SNP13, the question arises as to the importance of the common mutations in CD. Interestingly, c.-59G>A occurs within the same haplotype block as the CD-associated CARD15 polymorphisms. The precise haplotypes have to be defined. Therefore, we are currently identifying the haplotype structure in

relation to the promoter polymorphism. It was shown that CARD15 with the SNP13 mutation is functionally unable to detect bacterial peptides, resulting in less NFκB activation (Chamaillard *et al.* 2003; Inohara *et al.* 2003; Inohara *et al.* 2005). We here showed that a promoter polymorphism c.-59G>A influences the expression of CARD15 in response to TNF. Perhaps it is the combination of SNP8, SNP12 or SNP13 with the promoter polymorphism that causes susceptibility to CD.

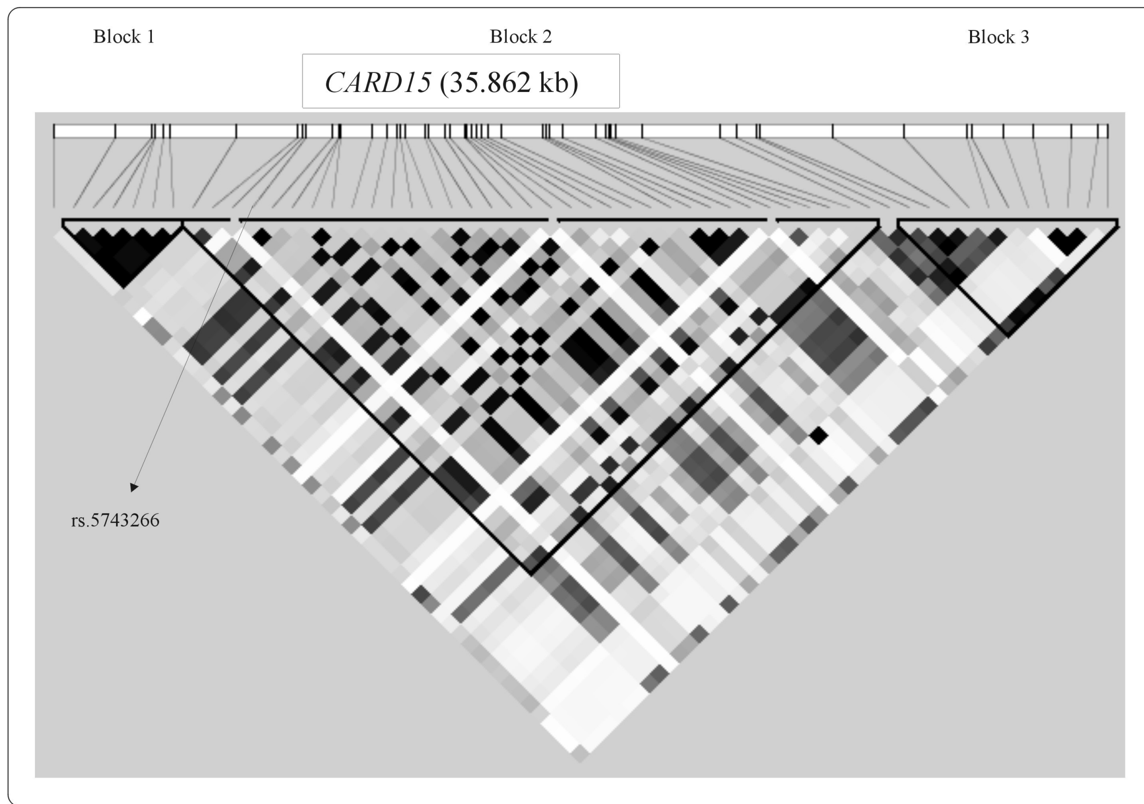


Figure 3 LD across *CARD15* and flanking genomic region, based on public CEPH data. R^2 values for pairwise LD between each marker are represented (white: r^2=0, shades of grey: 0>r^2>1, black: r^2=1). The top of the figure shows distances between markers. Most markers within the *CARD15* gene are in strong LD.

In summary, we found a new functional promoter polymorphism in *CARD15*, but to fully understand its association to CD, detailed haplotype analysis in our patient population is currently ongoing.

ACKNOWLEDGEMENTS

We acknowledge the International HapMap Consortium (HapMap 2003) for the CEPH data, and the Protein expression and Purification core facility of VIB for providing recombinant TNF.

REFERENCES

Berrebi, D., R. Maudinas, J. P. Hugot*, et al.* (2003). "Card15 gene overexpression in mononuclear and epithelial cells of the inflamed Crohn's disease colon." Gut. **52**(6): 840-6.

Chamaillard, M., S. E. Girardin, J. Viala*, et al.* (2003). "Nods, Nalps and Naip: intracellular regulators of bacterial-induced inflammation." Cell Microbiol. **5**(9): 581-92.

Cuthbert, A. P., S. A. Fisher, M. M. Mirza*, et al.* (2002). "The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease." Gastroenterology. **122**(4): 867-74.

Fang, Z. H. and Z. C. Han (2006). "The transcription factor E2F: a crucial switch in the control of homeostasis and tumorigenesis." Histol Histopathol **21**(4): 403-13.

Hampe, J., H. Frenzel, M. M. Mirza*, et al.* (2002). "Evidence for a NOD2-independent susceptibility locus for inflammatory bowel disease on chromosome 16p." Proc Natl Acad Sci U S A. **99**(1): 321-6.

HapMap (2003). "The International HapMap Project." Nature. **426**(6968): 789-96.

Hugot, J. P., M. Chamaillard, H. Zouali*, et al.* (2001). "Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease." Nature **411**(6837): 599-603.

Hugot, J. P., P. Laurent-Puig, C. Gower-Rousseau*, et al.* (1996). "Mapping of a susceptibility locus for Crohn's disease on chromosome 16." Nature **379**(6568): 821-3.

Inohara, Chamaillard, C. McDonald*, et al.* (2005). "NOD-LRR proteins: role in host-microbial interactions and inflammatory disease." Annu Rev Biochem. **74**: 355-83.

Inohara, N., Y. Ogura, A. Fontalba*, et al.* (2003). "Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease." J Biol Chem. **278**(8): 5509-12.

Laukens, D., H. Peeters, D. Marichal*, et al.* (2005). "CARD15 gene polymorphisms in patients with spondyloarthropathies identify a specific phenotype previously related to Crohn's disease." Ann Rheum Dis. **64**(6): 930-935.

Ogura, Y., D. K. Bonen, N. Inohara*, et al.* (2001a). "A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease." Nature **411**(6837): 603-6.

Ogura, Y., N. Inohara, A. Benito*, et al.* (2001b). "Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB." J Biol Chem **276**(7): 4812-8.

Ogura, Y., S. Lala, W. Xin*, et al.* (2003). "Expression of NOD2 in Paneth cells: a possible link to Crohn's ileitis." Gut **52**(11): 1591-7.

Rosenstiel, P., M. Fantini, K. Brautigam*, et al.* (2003). "TNF-alpha and IFN-gamma regulate the expression of the NOD2 (CARD15) gene in human intestinal epithelial cells."

Gastroenterology. **124**(4): 1001-9.

Shaw, S. H., J. Hampe, R. White*, et al.* (2003). "Stratification by CARD15 variant genotype in a genome-wide search for inflammatory bowel disease susceptibility loci." Hum Genet. **113**(6): 514-21.

van Heel, D. A., B. M. Dechairo, G. Dawson*, et al.* (2003). "The IBD6 Crohn's disease locus demonstrates complex interactions with CARD15 and IBD5 disease-associated variants." Hum Mol Genet. **12**(20): 2569-75.

Chapter 3

REDUCED METALLOTHIONEIN EXPRESSION IN COLONIC CROHN'S DISEASE: EVIDENCE FOR MTF1 AS A NEW DISEASE-MODIFYING GENE

Debby Laukens[1], Harald Peeters[2], Sara Bogaert[2], Anouk Waeytens[4], Bert Vander Cruyssen[3], Dirk Elewaut[3], Filip De Keyser[3], Herman Mielants[3], Claude Cuvelier[4], Eric M. Veys[3], Kenny Knecht[5], Paul Van Hummelen[6], Jurgen Del-Favero[7], Erik Remaut[1], Martine De Vos[2], Pieter Rottiers[1].

[1]Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium
[2]Department of Gastroenterology, Ghent University Hospital, Gent, Belgium
[3]Department of Rheumatology, Ghent University Hospital, Gent, Belgium
[4]Department of Pathology, Ghent University Hospital, Gent, Belgium
[5]Applied Maths BVBA, Sint-Martens-Latem, Belgium
[6]MicroArray Facility, VIB, Leuven, Belgium
[7]Genetic Service Facility, VIB, Antwerpen, Belgium

ABSTRACT

Background & aims: Crohn's disease (CD) is a multifactorial disorder involving both genetic and environmental factors. This study aimed at identifying and characterizing new candidate susceptibility genes for CD by integrating known disease loci with gene expression variation in non-inflamed colon biopsies of CD patients.
Methods: Sixteen CD patients and 11 controls were subjected to microarray analysis. Expression of metallothionein (MT) was analyzed in intestine and blood by quantitative PCR and immunohistochemistry. MT-knockdown HT29 cells were generated by small interfering RNA. *MT1M*, and its transcriptional regulator MRE-binding transcription factor 1 (*MTF1)*, were screened for mutations by sequencing.
Results: Eighteen differentially expressed baseline genes were identified. We focused on the

reduced expression of MT in intestine and blood of CD patients with colonic involvement. We showed that MT induction was not impaired in these patients. To model lowered MT expression in epithelial cells, we created MT-knockdown HT29 cells. These showed a reduced IL8 secretion in response to bacterial challenge, suggesting deficient inflammatory responses in the CD colon. No mutations were found in *MT1M* exons and its promoter region. A polymorphism in *MTF1* (IVS1-128A>T) was associated with disease location. Gene-gene interaction of *MTF1* and *CARD15* predisposes individuals to a high risk for ileal disease.
<u>*Conclusions*</u>: Combining microarray screening with genetic linkage data is an effective tool for identifying novel candidate susceptibility genes. We showed that deficient basal MT expression in CD patients with colonic involvement is genetically determined. The IVS1-128A>T polymorphism in *MTF1* is linked to disease location and serves as a new disease-modifying gene.

INTRODUCTION

The strong influence of genetic determinants in Crohn's disease (CD, MIM 266600) has been shown by familial clustering, and by the high concordance rate in monogenic twins. First-degree relatives of affected individuals show a 20 to 50-fold higher risk for developing CD. Moreover, affected siblings frequently develop the disease at similar ages, and concordance rates reach 80% for disease site, behaviour and presence of extra-intestinal manifestations (Peeters *et al.* 1996; Halfvarson *et al.* 2003). The mode of inheritance of CD is complex, and the number of genes predisposing to CD or modifying its course is currently unknown. However, it is not expected that a single risk gene is sufficient for disease development.
Two complementary approaches are used to identify determinants underlying genetically complex traits such as CD: candidate gene association studies and genetic linkage. In the candidate gene approach, genes that might be involved are screened for polymorphisms in a limited number of patients, and then frequencies in large populations of patients and controls are compared. In genetic linkage, candidate regions are identified by scanning the entire chromosome with microsatellite markers. These so-called disease loci can rarely be narrowed down to less than a few megabases. Nevertheless, if the candidate region contains a manageable number of genes, a candidate gene is chosen on the basis of a known property, e.g. function or tissue-specific expression. Unfortunately, mapping CD loci is hampered by phenotypic variability, genetic heterogeneity across populations, uncontrolled environmental influences, reduced penetrance, and limited statistical power in such studies. Nevertheless, several CD loci

have been corroborated by more than one independent study. The most frequently identified locus for CD is 16q12 (*IBD1*), which led to the mapping of the first gene to be firmly associated with CD, called *CARD15* (Hugot *et al.* 2001; Ogura *et al.* 2001). Today, there is evidence for more than 20 loci, but the identity of the causative genes remains largely unknown (Brant and Shugart 2004a). In addition, success in finding susceptibility genes has been limited by the modest effect of individual genes on the complex phenotype.

We applied a complementary approach to identify new potential susceptibility genes. Because much of the variation in gene expression is genetically transmitted (Cheung and Spielman 2002), we evaluated variation of gene expression in unaffected biopsies from CD patients, and combined these results with known linkage data. For this purpose, we looked at gene expression in biopsies taken from non-inflamed regions of the colon from CD patients and healthy controls. We selected genes that show variation in baseline expression and that are also located near a locus for CD. We focused on characterizing one of the candidate genes, metallothionein (MT), and the consequences of its altered expression. This led to the identification of a new disease-modifying gene, *MTF1*, located at 1p33 (*IBD7*).

MATERIALS AND METHODS

Patients and biopsies. CD was diagnosed according to clinical, endoscopic and histological criteria. Patients were classified according to the Vienna classification (Gasche *et al.* 2000). Disease location, however, was defined as the maximal spread of inflammation during the entire follow-up. We thus defined three subgroups: pure colonic involvement (C), both ileal and colonic involvement (IC), and ileal involvement only (I). For RNA extraction, 3 colon biopsies from each of 54 CD patients and 30 controls were collected during colonoscopy. All biopsies were taken from endoscopically normal regions of the sigmoid, immediately placed in RNAlater (Ambion, Cambridgeshire, UK) and stored at -80°C. For immunohistochemical analysis, colon biopsies were collected from 22 CD patients and 8 controls. Additional ileal biopsies were collected from 12 CD patients and 8 controls. Biopsy specimens were immersed in 4% formalin (Labonord, France).

RNA extraction. Total RNA was extracted from biopsies using the RNeasy Mini Kit (Qiagen, Westburg BV, The Netherlands) with on-column DNAse treatment (Qiagen). Needle homogenization was performed. The quality of RNA used for microarray analysis was checked on the Agilent 2100 Bioanalyzer (Agilent Technologies, Germany). Total RNA from 3 ml of

whole blood was extracted using the RNeasy Midi Kit (Qiagen) following lysis of red blood cells with erythrocyte lysis buffer (Qiagen). Concentration and purity (ratio OD_{260}/OD_{280} between 1.8 and 2.2) of the RNA used for quantitative real-time PCR was checked on a spectrophotometer (UV-1601, Shimadzu Benelux, Belgium).

Microarray hybridization, scanning and analysis. The construction of the focus microarray chip, containing 6,779 expressed sequence tags and specifically designed for the study of colonic gene expression, has been described (Laukens *et al.* 2006). Total RNA (5 µg) was amplified using a modified protocol for *in vitro* transcription (Puskas *et al.* 2002). Probe labelling, hybridization, washing and scanning were carried out at the MicroArray Facility (MAF, Leuven, Belgium) of the Flanders Interuniversity Institute for Biotechnology (VIB) as described on http://www.microarrays.be/service.htm. Images were analyzed with ArrayVision (Imaging Research Inc., Canada). Each hybridization was repeated in a dye swap experiment. Spot intensities were measured, corrected for local background, and those that exceeded the background by more than two standard deviations (SD) were included in the analysis. For each gene, ratios of red (Cy-5) to green (Cy-3) intensities (I) were calculated and normalized via a Lowess Fit of the $\log_2$ ratios [$\log_2$(ICy-5 / ICy-3)] over the $\log_2$ total intensity [$\log_2$(ICy-5 × ICy-3)].

A mixture of RNA from 5 CD patients, 5 non-CD inflammatory controls and 5 healthy controls served as reference RNA for comparison of the microarray datasets. This reference sample provides a positive hybridization signal at each probe element on the microarray, which is essential when calculating and comparing fluorescence ratios. The data were imported into GeneMaths® XT (Applied Maths, Belgium). Weighted mean ratios and their corresponding error (pixel SD) were calculated from the dye swap. Data were normalized over all arrays, and missing values were imputed using k-nearest neighbour algorithm (20 neighbours). GeneMaths® XT was used to perform all subsequent analyses.

Quantitative real-time PCR (qPCR). One µg of total RNA was converted to single stranded complementary DNA (cDNA) by reverse transcription (Superscript, Gibco, Invitrogen, Belgium) with oligo dT priming. One tenth of the cDNA was used in real-time quantification using the SYBR green kit (Eurogentec, Belgium) and 300 nM of each primer. A two-step program was run on the iCycler (BioRad Laboratories, Belgium). Cycling conditions were 95°C for 10 minutes and 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Melting curve analysis and direct

sequencing of amplicons on the ABI3700 analyzer (Perkin Elmer, Applied Biosystems, Belgium) confirmed primer specificities. All reactions were run in duplicate and normalized to glyceraldehyde phosphate dehydrogenase (GAPDH) levels. GAPDH was chosen after checking the expression stability of a set of housekeeping genes in biopsies of CD patients and controls using the Genorm software (Vandesompele *et al.* 2002). Primers were designed using the Beacon Designer software (PREMIER Biosoft International, USA). Sequences of all primer sets are listed in Table 1.

Table 1 Sequences of qPCR primer sets

Gene symbol	Reference sequence	Forward primer	Reverse primer
GAPDH	NM_002046	TGC ACC ACC AAC TGC TTA GC	GGC ATG GAC TGT GGT CAT GAG
MT1A	NM_005946	GCA AAG GGG CAT CAG AGA AGT G	AAA TAC AGT AAA TGG GTC AGG GTT G
MT1B	NM_005947	AAG TGC TGC TGC TCT TGC TG	TGG TTG CTC TAT TTA TGT CTG GGA G
MT1E	NM_175617	GCA TCC TCT GGG TCT GGG TTC	AAC AGC AGC CTG GGG AAG AAG
MT1F	NM_005949	GCG ACT GAT GCC AGG ACA AC	CAC AGG AAA AGG AAT GTA GCA AAT G
MT1G	NM_005950	GCA AAG GGG CAT CGG AGA AGT G	AAG GGA ATG TAG CAA AGG GGT CAA G
MT1J	AF348994	GCT GTG CCT GAT GTG GGA AC	AAA TGC AGC AAA TGG CTC AGT ATT G
MT1M	NM_176870	CTG CAA AGG GAC GTT GGA GAA C	CAG CAA ATG GCT CAG TAT CGT ATT
MT2A	NM_005953	AAA GGG GCG TCG GAC AAG TG	GAA TAT AGC AAA CGG TCA CGG TCA G
MTF1	NM_005955	TAA GAC TCA ATT GAT TCA GGG ACG AGA	CAG TTG TGA GAA ATG AAA ACG TAA TGA

Immunohistochemistry. Colon and ileum biopsies fixed in formalin were embedded in paraffin according to standard procedures. Sections were rehydrated by serial immersion in xylene and ethanol. Immunostaining was performed on a NexES IHC automated staining system (Ventana Medical Systems, France) using a 1:10 dilution of a mouse monoclonal anti-metallothionein antibody (clone E9, Zymed Laboratories, Sanbio, The Netherlands). An isotype-specific irrelevant antibody (X0931, DakoCytomation, Belgium) was used to control for non-specific binding of the primary antibody. Slides were blindly scored by two pathologists. Semi-quantitative scoring was based on staining intensity and the number of positive cells.

Induction assays. The mononuclear cell fraction (PBMC) was isolated from 30 ml of venous blood by density gradient centrifugation in Ficoll-Pacque (Amersham Biosciences, The Netherlands). PBMCs were seeded in 6-well plates at 10^6 cells per well in RPMI medium (Gibco, Invitrogen) supplemented with 10% FCS. The next day, they were stimulated with 200 μM $ZnSO_4$ (Sigma, Belgium), 2 μM dexamethasone (Sigma), 10 ng/ml recombinant human IL6 (Peprotech, Campro Scientific, The Netherlands), 1000 U/ml human recombinant

TNF (produced in-house) or 200 U/ml human recombinant IFNγ (Biosource Europe, Belgium). The cells were lysed in 350 µl RLT buffer (Qiagen) after 3, 6 or 24 hours. To induce oxidative stress, PBMCs were incubated with 100 µM H_2O_2 (Sigma), which was washed off with PBS after 10 minutes. Cells were lysed after 2, 4 or 6 hours. Total RNA was isolated using the Rneasy Mini Kit (Qiagen), and converted to cDNA for subsequent qPCR. Normalized Ct values were corrected for background at each time point (e.g. $dCt_{induced,\ t1} - dCt_{un\text{-}induced,\ t1}$), and plotted against time (0, 3, 6 and 24 hours, or 0, 2, 4 and 6 hours for H_2O_2). Induction was calculated as the area under the curve (AUC) for the three time points (AUCs are shaped as triangles and rectangles, and their summation or integration gives the same result):

$$AUC = \int_0^3 (m_1.x)dx + \int_3^6 (m_2.x + b)dx + \int_6^{24} (m_3.x + c)dx,$$

where m_1 = the slope of the 0-3 line, m_2 = the slope of the 3-6 line, m_3 = the slope of the 6-24 line, b = the y-intercept of the 3-6 line, and c = the y-intercept of the 6-24 line.
HT29 human colon epithelial cells (American Type Culture Collection (ATTC) HTB38) and MT-knockdown cells (HT29MTkd) were seeded in 24-well plates at 10^6 cells per well in RPMI supplemented with 10% FCS. The following day, cells were infected in triplicate at multiplicities of infection (MOI) of 100 with *adherent-invasive Escherichia coli* strain LF82, isolated from a patient with CD, (a gift from A. Darfeuille-Michaud) (Darfeuille-Michaud *et al.* 2004) for 1 hour, followed by gentamycin treatment (100 µg/ml) for 1 hour. Supernatants were collected after 8, 24 and 48 hours of infection/induction.

Small interfering RNA (siRNA). The short hairpin RNA (shRNA) fused to an H1 promoter was synthesized by PCR on the pSUPER vector (Tronolab, Switzerland), using primers: 5'-CAA TCT CTT GAA **TTG CAC TTG CAG GAG CCG** GGG GGA TCT GTG GTC TCA TAC AGA ACT TAT AA-3' and 5'-CCA TCG ATT TCC AAA AA**C CGG CTC CTG CAA GTG CAA** TCT CTT GAA TTG C-3'. The sequence in bold is a 19-mer that specifically targets MT, and is present in MT1B, MT1E, MT1H, MT1J and MT1M. The PCR fragment was cloned in pLVTH-siGFP (Tronolab, Switzerland). To produce virus for delivery of the shRNA, HEK293T cells (ATCC CRL1573) were transfected by the calcium phosphate method with 3 µg pCMV-d8.91 (Tronolab, Switzerland), 1.5 µg PMDG2 (Tronolab, Switzerland) and 1.5 µg of the shRNA construct. After 48 hours, supernatant containing viral particles was harvested, and put through a 0.45 µm filter. One day before viral transduction, HT29 cells were seeded at $5x10^4$ cells per well in a 24-well plate. They were overlaid with viral supernatant, and centrifuged for 1 hour at 32°C. This procedure was repeated with 72-hour viral supernatant. Cells expressing high

levels of GFP were subsequently sorted with the EPICS altra cell sorter (Beckman Coulter, The Netherlands).

Cytometric Bead Array immunoassay. Concentrations of IL1β, IL6, IL8, IL10, IL12p70 and TNF in the culture supernatants were measured simultaneously using Cytometric Bead Array (CBA) (Human Inflammation Kit, Pharmingen, Becton Dickinson, Belgium) according to the manufacturer's instructions. Concentrations of the cytokines were obtained by comparing the mean fluorescence intensity of the samples with that of the corresponding standard curves. Flow cytometric analysis was performed using a BD FACS scan. Data acquisition and analysis was done using BD CBA software.

IL8 ELISA. IL8 concentrations in supernatants were assessed in dilution series by sandwich ELISA. Microtiter plates (Nunc, USA) were coated with 2 μg/ml anti-IL8 antibody (Pharmingen) in PBS for 2 hours at RT, and non-specific binding sites were blocked with 0.1% casein-PBS at 4°C overnight. Dilution series of the supernatant (1:2) were incubated for 2 hours at RT in 0.1% casein-PBS, and detected with 1:1000 biotinylated anti-IL8 antibody (Pharmingen) and 1:500 streptavidin-HRP (Pharmingen) in 0.1% casein-PBS for 1 hour. Substrate was added and the reaction was stopped after 30 minutes with 1 M H_2SO_4. Plates were read at 450 and 595 nm in a microplate reader (BioRad).

Mutation screening of *MT1M* and *MTF1*. Genomic DNA was extracted from whole blood using the Qiagen blood and cell culture DNA kit. The sequence encompassing the *MT1M* gene and ~600 bases upstream of the transcription start site was PCR-amplified in 25 randomly selected CD patients and 6 controls (forward primer: 5'-GAG GCC GAC CAG TGT TC-3', reverse primer: 5'-AGT CTC TGG ATG AAA ATG TGA G-3'). The PCR product was sequenced with BigDye Terminator v3.1 Cycle Sequencing kit and analyzed on an ABI3700 analyzer (Applied Biosystems) with the former and the following primers: exon 1: 5'-CCC AGC CCA GCC CAG GAC CG-3'; exon 2: 5'-TCA CTG CCC ACT GCG TTT TTC TC-3'; exon 3: 5'-CAA GTC TAC TGC TAC CTC TC-3'.

Mutation analysis of *MTF1* was performed in 95 randomly selected CD patients using flanking intronic primers for all exons and overlapping primers for 1 kb of the upstream regulatory region. PCR primers were designed using the SNPbox software (Weckx *et al.* 2005a). A total of 28 primer sets were used (primer sequences are available on request). PCR fragments were

sequenced with BigDye Terminator v3.1 Cycle Sequencing kit (Perkin-Elmer, Applied Biosystems), and analyzed on an ABI3730 DNA analyzer (Applied Biosystems). The sequencing trace files were analyzed for the presence of variants using novoSNP (Weckx *et al.* 2005b).

***CARD15* and *MTF1* genotyping.** The 3 CD-associated *CARD15* polymorphisms R702W (SNP8), G908R (SNP12) and 1007fs (SNP13), as well as *MTF1* IVS1-128A>T, were genotyped in a cohort of 222 randomly selected CD patients and 63 controls using RFLP-PCR as previously described (Laukens *et al.* 2005). *MTF1* IVS1-128A>T genotyping was also done by RFLP-PCR. A fragment containing the polymorphism was amplified with the following primers: forward 5'-TAA GAC TCA ATT GAT TCA GGG ACG AGA GC-3', reverse 5'-CAG TTG TGA GAA ATG AAA ACG TAA TGA C-3'. The PCR product was precipitated with SeeDNA (Amersham Biosciences), digested with *Dra*I, and restriction fragments were analyzed on a 2% agarose gel. The presence of the IVS1-128T allele abolishes a *Dra*I site, resulting in a band of 271 bp, instead of 251 bp in the presence of the IVS1-128A allele.

Statistical analysis. Statistical analyses were performed using SPSS software (SPSS inc., USA). Differences between groups in immunohistochemical scores and induction AUCs were calculated using the Mann-Whitney U test. Correlations between metric data were calculated using Pearson's Rho, and those between rank data with Spearman's Rho. Rank data for histological scores were 0-1 = 1, 1-2 = 2, 2-3 = 3 and for RNA levels <1,000 = 1, 1,000-10,000 = 2, >10,000 = 3. Odds ratios and their 95% confidence intervals (CI) were calculated with Pearson's χ^2 test. Logistic regression was used to evaluate independence and interaction between parameters. Significance of differences in ELISA IL8 concentrations between time-series were calculated using a general linear model, error bars are calculated from 4 dilutions. Two-tailed probabilities were calculated, and *P*-values of less than 0.05 were considered statistically significant.

RESULTS

Identification of potential candidate genes for CD. Gene expression in normal colon biopsies of 16 CD patients and 11 healthy controls (Table 2) was analyzed on a focus microarray chip containing 6,779 expressed sequence tags. We aimed at identifying potential candidate susceptibility genes for subsequent validation and mutation screening. This process was

performed in two steps. First, we characterized genes that were differentially expressed between CD patients and controls using an independent t-test (*P*<0.01, *N*=240). We then selected the genes that were located near a chromosomal region that has been linked to CD in at least two independent studies (Table 3). We also included the loci that were found in a study performed on a Flemish cohort (Vermeire *et al.* 2004a), because the patients in our study live in this area of Belgium. We employed an arbitrary distance of 5 megabases around markers that showed the highest linkage. The exact chromosomal locations of the markers and genes were explored using the genome browser at http://genome.ucsc.edu. In this way, 18 genes were selected (Table 4), and all clones of the respective genes were sequence verified and annotated correctly.

Table 2 Clinical characteristics of CD patients analyzed by microarrays

Patient	Sex	Familial CD	Age of onset	Behaviour	Disease location	*CARD15* status	Operations	Medication
1	F	no	A1	B1	IC	wt	no	immunosuppressives
2	M	no	A1	B3	IC	ND	no	no
3	M	no	A2	B1	IC	mutant	no	5-ASA
4	F	yes	A1	B3	IC	wt	resection	corticosteroids
5	M	no	A1	B3	IC	mutant	fistula	5-ASA
6	F	no	A1	B2	C	mutant	resection	Remicade
7	M	no	A2	B2	I	wt	resection	5-ASA
8	F	no	A1	B3	IC	wt	resection	immunosuppressives
9	F	yes	A1	B3	IC	wt	fistula	immunosuppressives
10	F	yes	A2	B2	I	wt	resection	immunosuppressives
11	F	no	A1	B3	IC	mutant	fistula	immunosuppressives
12	M	yes	A1	B2	I	wt	no	5-ASA
13	F	no	A1	B3	C	wt	resection	immunosuppressives
14	F	no	A1	B1	IC	wt	no	5-ASA
15	F	no	A1	B3	IC	wt	fistula, resection	corticosteroids
16	F	no	A1	B3	IC	wt	fistel	5-ASA

A1: <40 years; A2: ≥ 40 years; B1: non-stricturing, non-penetrating; B2: stricturing; B3: penetrating; disease location is defined as maximal extension of inflammation during total follow-up of patients with I: ileal involvement only, C: colonic involvement only, IC: ileal and colonic involvement; mutant *CARD15*: carriage of at least one mutant allele for SNP8, SNP12 or SNP13 (according to Hugot *et al.* 2001); wt *CARD15*: not carrying a mutant allele for SNP8, SNP12 or SNP13; 5-ASA: 5-aminosalicylates

Metallothioneins are down-regulated in colon, ileum and whole blood of CD patients with colonic involvement.

Metallothionein RNA expression. In the microarray screen we found two closely related metallothionein (MT) transcripts, MT1F and MT1M, that were down-regulated in CD patients. Although other genes may show lower *P*-values, we chose to focus on these genes, because they are located near the *IBD1* locus, which has been linked to CD in almost every linkage analysis,

Table 3 Differentially expressed genes in colon biopsies of CD patients located near a locus for CD

Unigene	Gene symbol	Gene description	Chromosomal location	Expression in CD	*P*-value
Hs.486246	PHTF1	putative homeodomain transcription factor 1	1p13	↑	0.000099
Hs.77955	MEF2D	MADS box transcription enhancer factor 2, polypeptide D	1q22	↓	0.0035
Hs.106674	BAP1	BRCA1 associated protein	3p21.31-p21.2	↓	0.0017
Hs.438691	GMPPB	GDP-mannose pyrophosphorylase B	3p21.31	↓	0.0013
Hs.302047	PLCL3	phospholipase C-like 3	3q25.31	↓	0.0048
Hs.318567	NDRG1	N-myc downstream regulated gene 1	8q24.3	↓	0.00038
Hs.84072	TM4SF3	transmembrane 4 superfamily member 3	12q12.1	↑	0.0046
Hs.85951	XPOT	exportin, tRNA (nuclear export receptor for tRNAs)	12q14.1	↑	0.0021
Hs.419776	NAP1L1	nucleosome assembly protein 1-like 1	12q21.1	↓	0.0088
Hs.159481	GALGT	UDP-N-acetyl-alpha-D-galactosamine	12q13.3	↓	0.00054
Hs.438737	MT1F	metallothionein 1F	16q12.2	↓	0.001
Hs.188518	MT1M	metallothionein 1M	16q12.2	↓	0.00058
Hs.100914	CEP192	Centrosomal protein 192kDa	18p11.21	↑	0.000042
Hs.512640	PRKCSH	protein kinase C substrate 80K-H	19p13.2	↓	0.00042
Hs.134074	SLC35E1	solute carrier family 35, member E1	19p13.11	↑	0.006
Hs.437	TCF15	transcription factor 15 (basic helix-loop-helix)	20p13	↓	0.0024
Hs.102336	ARHGAP8	Rho GTPase activating protein 8	22q13.31	↓	0.0093
Hs.28491	SAT	spermidine/spermine N1-acetyltransferase	Xp22.1	↑	0.0035

↑: up-regulated in CD; ↓: down-regulated in CD

and which contains *CARD15*. Metallothioneins are involved in a variety of metal scavenging processes, and they possess both bactericidal and anti-inflammatory activities (Zangger *et al.* 2001; Itoh *et al.* 2005). In humans, at least 18 MT gene isoforms exist (*MT1A-MT1X*, *MT2A*, *MT3* and *MT*4), but only *MT1* and *MT2* are inducible. We performed a detailed expression study on different MT isoforms using qPCR on an independent population of 38 CD patients and 19 controls. Because of the high sequence homology between MT isoforms, we first determined whether both MT1M and MT1F were specifically down-regulated. To this end, we designed primers for MT1M and MT1F, as well as for MT1E, MT1J and MT2A (Table 1).

Primer specificities were confirmed by the presence of a single melting peak after denaturation of the amplicons, and by direct sequencing. The expression of MT1M in CD patients with colonic involvement (subgroups C and IC) was significantly less than in patients with pure ileal involvement (P=0.0244, Figure 1) and in controls (P=0.0076). Furthermore, the expression of MT1M in all biopsies tested was significantly correlated with expression of the other isoforms tested (Table 5). This indicates that at least these MT isoforms share the same regulatory machinery for basal expression in the colon. The highest expression in colon biopsies was observed for MT1E, followed by MT2A>MT1J>MT1F>MT1M.

Table 4 Correlations of MT isoform expression in colon biopsies of CD patients and controls

	MT1E	MT1F	MT1J	MT1M	MT2A
MT1E Pearson's Rho	1	0.631	0.496	0.500	0.499
P-value		<0.01	<0.01	<0.01	0.058
N		65	55	66	15
MT1F Pearson's Rho		1	0.589	0.378	0.784
P-value			<0.01	<0.01	<0.01
N			55	62	15
MT1J Pearson's Rho			1	0.609	0.805
P-value				<0.01	<0.01
N				53	13
MT1K Pearson's Rho				1	0.615
P-value					<0.05
N					15
MT12A Pearson's Rho					1
P-value					
N					

Basal transcript levels of MT1F were measured in whole blood of 15 CD patients and 13 controls, and were found to be down-regulated in CD patients with colonic involvement (*P*=0.026, Figure 2). MT1M levels were too low for reliable quantitative measurement. None of the patients and none of the controls had elevated C-reactive protein levels at the time of blood sampling. Therefore, changes in MT expression due to systemic inflammation is not likely (Vermeire *et al.* 2004b).

Metallothionein protein expression. MT protein expression was evaluated semi-quantitatively in colon and ileum biopsies of CD patients and controls using immunohistochemical staining with a mouse monoclonal anti-metallothionein antibody (cloneE9, Zymed Laboratories). This antibody however, cannot distinguish between MT isoforms, because all of them have the epitope it recognizes. Overall, the expression of MT in colon was significantly lower than in the ileum (*P*=0.042). No conclusions could be drawn about MT expression in the colon, because the scores rarely reached 2. Compared to controls, a significant decrease in MT protein expression was found in CD patients with colonic involvement (subgroups C and IC) in ileum biopsies, again indicating a possible fundamental defect in MT expression in these patients (data not shown). A correlation was found between RNA expression as assessed by qPCR and the immunohistochemical score of protein expression (Spearman's Rho: 0.826, *P*=0.001, *N*=12).

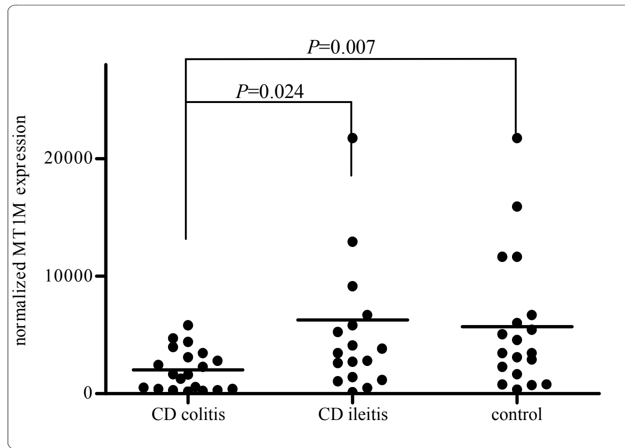


Figure 1 MT1M expression is down-regulated in colon biopsies of CD patients with colonic involvement. CD colitis: subgroups IC + C (*N*=20), CD ileitis: subgroup I (*N*=18), control (*N*=19).

This finding is interesting because in each individual, the biopsies used for RNA extraction and those used for staining were not obtained at the same time, again indicating a stable inherent basal expression of MT. Epithelial cells were primarily positive for MT staining. The strongest expression was found at the base of the villi, within paneth cells, and in rapidly proliferating epithelial cells in the crypts (Figure 3).

Metallothionein induction in response to oxidative stress is higher in CD patients. Because MT expression is decreased in the colon, ileum and whole blood of CD patients with colonic involvement, we examined whether this was due to impaired induction. We stimulated PBMC isolated from 10 CD patients (subgroups C: 2, IC: 4, I: 4) and 9 controls with known MT inducers (10 ng/ml IL6, 1000 U/ml TNF, 200 μM $ZnSO_4$, 200 U/ml IFNγ and 2 μM dexamethasone), and measured MT1M expression by qPCR. No difference in induction was found between CD patients and controls. However, induction following exposure to 100 μM H_2O_2 for 10 minutes was significantly higher in CD patients, independently of the disease location (*P*=0.0015, Figure 4).

Colonic epithelial cells defective in metallothionein expression secrete less IL8 in response to LPS or *adherent-invasive E. coli*. Because we observed MT expression mainly in epithelial cells, we investigated whether a low basal expression of MT in epithelial cells has consequences upon challenge with bacteria. We used small interfering RNA (siRNA) to create HT29 colonic epithelial cells deficient in MT expression (HT29MTkd). Although the siRNA sequence was targeted against MT1B, MT1E, MT1H, MT1J and MT1M, the RNA expression of most MT isoforms was lowered approximately twofold as assessed by qPCR (Figure 5). Diminished

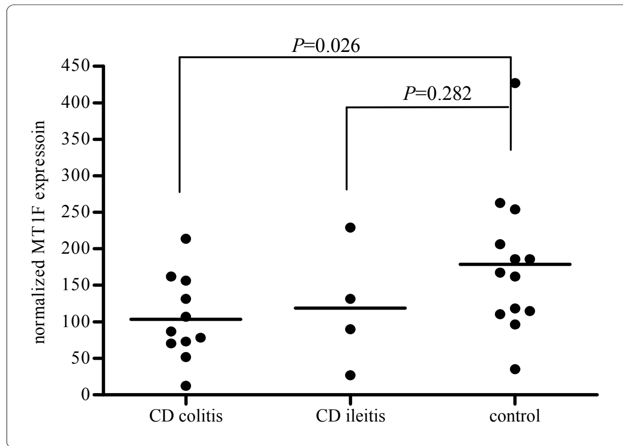


Figure 2 MT1F expression in blood is down-regulated in CD patients with colonic involvement. CD colitis: subgroups IC + C (*N*=11), CD ileitis: subgroup I (*N*=4), control (*N*=13).

protein expression was shown by immunofluorescence using the monoclonal anti-metallothionein antibody (data not shown). Cytokine secretion in response to stimulation with *adherent-invasive E. Coli* (*AIEC LF82*) was tested with the Human inflammation kit (Becton Dickinson). We could only measure IL8 secretion after bacterial challenge with this technique. The secretion of IL8 was significantly lower in HT29MTkd cells than in normal HT29 cells (Figure 6).

Screening for mutations in the *MT1M* gene. We screened the complete *MT1M* gene (~1.4 kb) and its promoter region (~600 bases upstream of the transcriptional start site), using NM_176870 as reference sequence, in 25 CD patients and 6 controls. Two intronic polymorphisms, IVS1-366C>T and IVS2-49C>T, and one 3' untranslated region (3'UTR) polymorphism, c.286A>G, were found. The absence of promoter or coding mutations in *MT1M* that would cause its down-regulation was not too surprising, since the down-regulation was not confined to MT1M. Better candidates for mutation screening are probably upstream factors regulating all MT isoforms, and so we proceeded in that direction.

A polymorphism in the first intron of the *MTF1* transcription factor gene is associated with ileal disease. One of the best-characterized transcription factors regulating MT expression is MRE-binding transcription factor 1 or *MTF1*, an essential zinc finger protein that binds to specific DNA motifs termed metal-response elements (MRE). Furthermore, this protein is responsible for both basal and inducible expression of MT (Samson and Gedamu 1998), and the gene is located at 1p33 (*IBD7*), a locus that has been linked to CD in a Flemish population

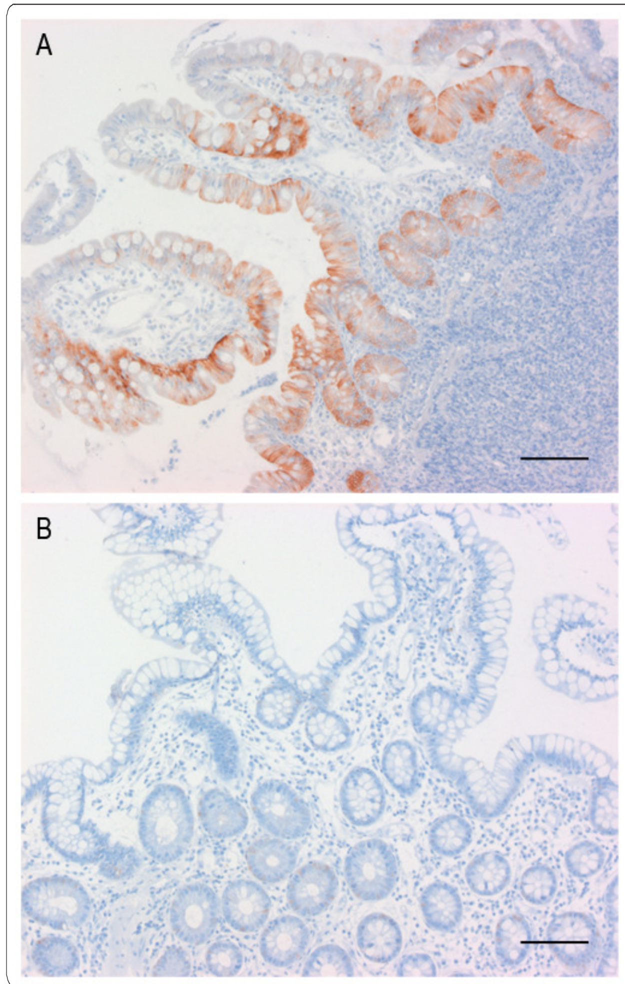


Figure 3 Metallothionein staining of representative samples of non-inflamed ileal biopsies of (A) a control and (B) a CD patient. Control ileum shows strong immunoreactivity for MT in epithelial cells, with stronger reactivity in the proliferative cells of the crypts and at the base of the villi. In CD patients, MT expression is significantly decreased.
Bar=100 µm

(Vermeire *et al.* 2004a). We show that expression of the transcription factor MTF1 in blood of CD patients and controls correlates with the expression of MT1F (Pearson's Rho: 0.572, P=0.001, N=20). All 11 exons and the promoter region (~1000 bp upstream of the transcription start site) were screened for mutations in 95 CD patients, using NM_005955 as a reference sequence. Two missense mutations were found: c.198C>G (Asp63Glu) in 4 patients, and c.1253G>A (Glu385Lys) in 2 patients. The functional relevance of these mutations still needs to be investigated. Moreover, a potential polymorphism at the splice site junction between exons 8 and 9, c.1270A>G, was frequently found. Yet no alternative splice variant could be detected in cDNA from blood of patients with different genotypes (data not shown). Here, we focused on a polymorphism in the first intron of *MTF1*, IVS1-128A>T, because of its potential influence on gene expression (Kleinjan and van Heyningen 2005). No difference in frequency between a cohort of 222 CD patients and 63 controls was detected (data not shown). However, genotype-phenotype analysis revealed that IVS1-128A>T had considerable influence on the location of

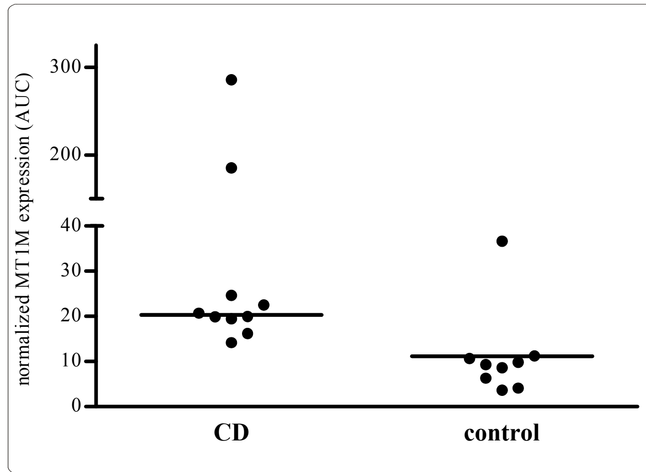


Figure 4 MT1M RNA induction in response to oxidative stress is higher in CD patients (*N*=10) than in controls (*N*=9), irrespective of the disease location (*P*=0.0015).

disease: the IVS1-128T allele was associated with ileal disease (subgroups I and IC). In the presence of the IVS1-128T allele, 139/166 patients (84%) had ileal disease, compared to 38/56 (68%) of those with the AA genotype (OR: 2.4, CI: 1.216-4.891, *P*=0.011). Because we observed that *CARD15* mutations were also highly correlated with ileal disease (subgroups I and IC, *P*<0.0001), which has also been observed by others (Ahmad *et al.* 2002; Cuthbert *et al.* 2002; Lesage *et al.* 2002), we investigated whether the combined presence of *CARD15* and *MTF1* risk genotypes had a substantially higher impact on disease location than either of them alone. The odds ratio for developing ileal disease (subgroups I and IC) in *CARD15* mutation carriers is 4.2 (CI: 1.980-9.098) and in IVS-128T allele carriers 2.4 (CI: 1.216-4.891). Moreover, logistic regression revealed significant interaction between the two genetic markers: carriage of an IVS-128T allele increased the risk of ileal disease by 12% in *CARD15* wild type patients (from 60 to 72%, Figure 7), and in carriers of *CARD15* mutations by 23% (from 74 to 97%). This means that both genes contribute to the location of inflammation, and mutually interact.

The IVS-128A>T polymorphism is located in the first intron of *MTF1*, which is located within the 5'UTR. It is not contained within a CpG island. A search for transcription factor binding sites using the transcription factor database (TRANSFAC) revealed a potential binding site for GATA binding protein 4 (GATA4).

DISCUSSION

Identification of susceptibility genes, their interaction, and their relationship to specific clinical manifestations is an important step in understanding CD and developing improved clinical

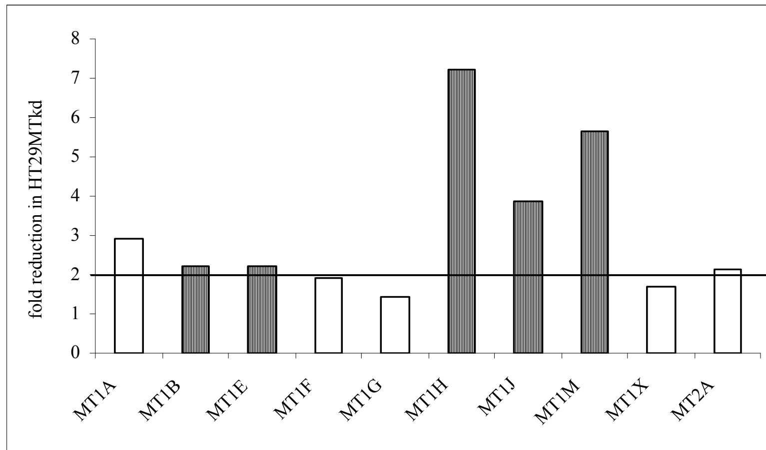


Figure 5. The RNA expression of different MT isoforms in HT29MTkd cells is down-regulated in comparison to HT29. The solid line represents a twofold change in expression. The short hairpin RNA was targeted at MT1B, MT1E, MT1H, MT1J and MT1M (black bars), but the other isoforms tested (MT1A, MT1F, MT1G, MT1X and MT2A) also show reduced expression.

management or therapy. We used an alternative, hypothesis-independent strategy to identify novel candidate genes. Transcriptome analysis of non-inflamed colon biopsies of CD patients and controls revealed 18 novel potential candidate genes localized to CD loci.

Many genome-wide linkage studies of CD have been performed (Brant and Shugart 2004b). This led to the identification of the first gene strongly associated with CD, *CARD15* (Hugot *et al.* 2001; Ogura *et al.* 2001). Nevertheless, despite its strong association with CD, *CARD15* is difficult to relate causally to CD. Polymorphisms show very limited penetrance, occurring in approximately 40% of patients (predominantly those with ileal disease), as well as in 15% of healthy individuals. Furthermore, linkage to chromosome 16 was still observed in CD patients not carrying common *CARD15* mutations (Hampe *et al.* 2002; Shaw *et al.* 2003; van Heel *et al.* 2003). An alternative explanation is that polymorphisms in *CARD15* are not in themselves causal, but modify the immune response in inflammatory lesions elicited by some other mechanism. An interesting theory based on evolutionary benefit was suggested by Hugot and co-workers (Hugot *et al.* 2003). They stated that a mutated CARD15 protein would have been beneficial during the outbreak of the plague in Europe, but it somehow represents a disadvantage in the pathogenesis of CD. The identification of *CARD15* and the insights into its role in innate immunity and CD pathology (Eckmann and Karin 2005) highlight the importance of mapping susceptibility genes.

Two of the 18 potential candidate genes we identified belong to the family of closely related metallothioneins (MT). They are located in tandem within the *IBD1* locus and arose by non-processed gene duplications (Karin *et al.* 1984). Furthermore, they are involved in protecting cells against toxic levels of metal ions, radicals and bacterial infections. Four MT classes exist in humans, but only MT1 and MT2 isoforms are inducible by cytokines, hormones, metals and

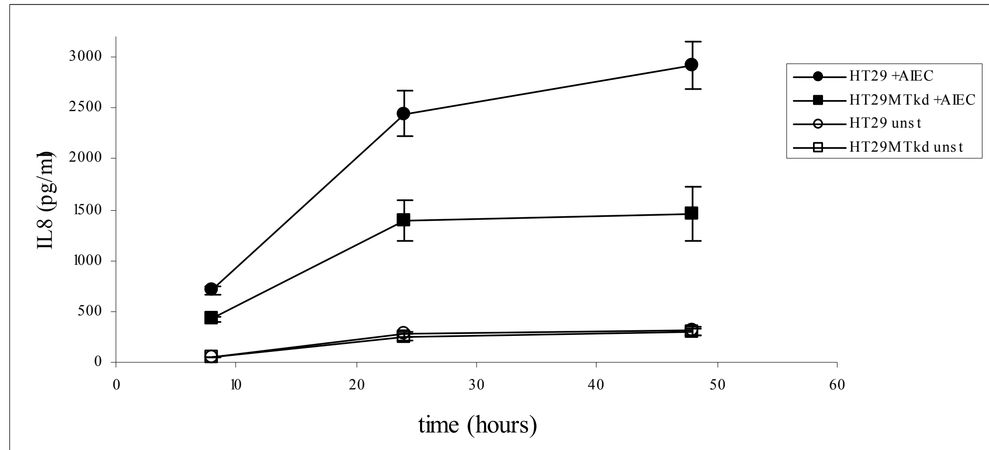


Figure 6. IL8 secretion in response to bacterial challenge is lower in MT-knockdown cells (HT29MTkd) than in HT29 cells. +AIEC: stimulated with *adherent-invasive E. coli* LF82 MOI 100; unst: unstimulated ($P<0.05$).

stress in general (Haq *et al.* 2003). We showed that MT was significantly down-regulated in CD patients with colonic involvement as compared to healthy controls. In addition, detailed expression analysis of different MT isoforms revealed that their expression is highly correlated at basal level. Low RNA expression levels were linked to reduced protein expression. Furthermore, we provide evidence that a low expression level of MT in CD patients with colonic involvement results from a genetic predisposition rather than from an early inflammation event. Indeed, the down-regulation was not only found in colon but also in ileum biopsies, both sampled from non-inflamed areas, as well as in whole blood samples regardless of C-reactive protein levels. Remarkably, MT RNA levels correlated with protein expression in biopsies that were not time matched, contributing to the idea that MT expression is stable and potentially inherited. Two studies have previously shown an up-regulation of MT in CD (Bruwer *et al.* 2001; Dooley *et al.* 2004), while we and others found a down-regulation (Clarkson *et al.* 1985; Elmes *et al.* 1986; Ioachim *et al.* 2003; Kruidenier *et al.* 2003). Moreover, MT quantification in CD based on radioimmunoassay (Mulder *et al.* 1991), microarray (Lawrance *et al.* 2001) and silver-saturation assay (Sturniolo *et al.* 1998) also reported a down-regulation.

We further demonstrated that although basal MT levels were reduced in CD patients with colonic involvement, they are still inducible by conventional stimuli in peripheral blood. Unexpectedly, MT induction with oxidative stress was higher in CD patients, independent of disease location. Relative overreaction in CD patients might reflect the fact that basal low levels of MT result in more free hydroxide radicals that can activate the MT transcription factor, MTF1, resulting in the overall increased transcription of MT (Zhang *et al.* 2003). Alternatively,

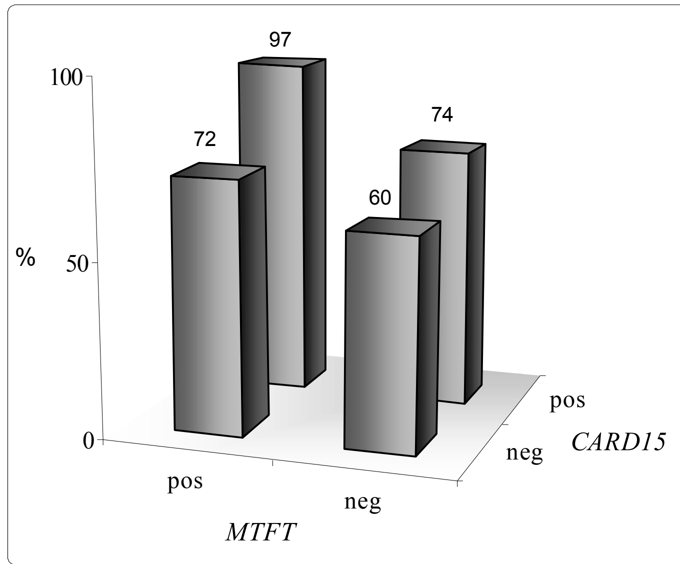


Figure 7 The presence of the *MTF1* IVS-128T allele and *CARD15* mutations contribute to ileal disease. Percentages represent number of patients positive or negative for *CARD15* mutations that are positive or negative for the IVS-128T allele. Logistic regression revealed a gene-gene interaction between *CARD15* and *MTF1*.

the higher basal MT level in controls may not permit further induction, because a certain threshold is reached.

In order to clarify the potential functional consequence of reduced basal levels, we generated MT-knockdown colonic epithelial HT29 cells using small interfering RNA. We report that, in response to bacterial challenge with a CD-associated *E. coli* strain (Darfeuille-Michaud *et al.* 2004), reduced expression of MT in epithelial cells correlated with reduced IL8 secretion. This observation is in accordance with the recent finding that IL8 up-regulation is impaired in response to acute trauma to colon and skin of CD patients (Marks *et al.* 2006). IL8 is a potent chemoattractant and activator of neutrophils, which respond rapidly to different types of infections, and play an essential role in the inflammatory response. They release antimicrobial peptides as well as reactive oxygen intermediates that may cause tissue damage. We hypothesize that the intestinal epithelial cell lining of CD patients with colonic involvement expresses less MT and initially secretes less IL8 in response to bacteria. This might secure an exaggerated secondary, compensatory immune response.

It is unlikely that polymorphisms in the MT genes themselves would be responsible for their low expression, since we showed that they are down-regulated as a group. Yet, in our knockdown model for MT, we see a similar decrease in expression of MT isoforms that were not targeted with the short hairpin RNA. Thus, the possibility of a positive feedback loop for MT expression via its own transcriptional regulator cannot be ruled out (Kimura *et al.* 2002). We could not find coding or promoter polymorphisms in *MT1M*, but other isoforms need to be screened. Alternatively, epigenetic mechanisms such as DNA methylation might be responsible for the

down-regulation of MT genes in CD patients (Smith *et al.* 2005).

In searching for mutations in *MTF1* (*IBD7*), a transcription factor that regulates basal expression of MT, we focused on a polymorphism within the first intron, IVS1-128A>T, because regulating regions are frequently found there. This polymorphism predisposes to inflammation at specific sites in the intestine. Disease location is a phenotypic characteristic that remains stable over time. Similar to *CARD15*, the association of *MTF1* with ileal disease offers a new candidate disease-modifying gene, rather than a disease-predisposing gene (Gasche *et al.* 2003). In addition, the two genes interact to bring about ileal disease. If we regard pure colonic CD on one end of a continuum with ulcerative colitis (UC) on the other end, this polymorphism in UC patients is worthy of study.

The links of *CARD15* and *MTF1* with disease location might reflect the functions of these proteins at their sites of expression: CARD15 is expressed in paneth cells of the ileum, and the expression of MT in ileum is much higher than in colon. Both proteins are probably crucially involved in maintaining a low bacterial count in the ileum. A fully functional CARD15 or MTF1 might be crucial in this process. However, they might be disadvantageous in the colon, where bacterial load is high. In the colon, a weakly functional CARD15 can be beneficial, because inappropriately intense immune responses will not be elicited. Greater MT expression in the colon can also be beneficial because it can lead to efficient clearance of mucosal infection.

The association of the IVS1-128T allele in *MTF1* with low MT expression needs investigation. However, to fully understand the genetic contribution of IVS-128A>T in CD, the haplotype structure surrounding this polymorphism needs to be determined. The mutation causing altered transcription of MTF1 could be located elsewhere within the same haplotype block. MTF1 regulates basal expression of MT (Samson and Gedamu 1998). Therefore, a difference in MTF1 expression will be directly linked to altered MT expression. Indeed, we showed a correlation between MT1F and MTF1 levels in peripheral blood. The IVS-128T allele is located within a potential binding site for GATA4, a transcription factor that is expressed in the ileal enterocytes in the small intestine, but not in the colon (Boudreau *et al.* 2002). Therefore, a primary goal in this context is to study GATA4 binding to the *MTF1* gene.

Loss of MTF1 in mouse embryonic fibroblasts results in enhanced collagen deposition, which is an important complication of CD (Haroon *et al.* 2004). In these fibroblasts, transforming growth factor-beta is activated. This protein has potent anti-inflammatory properties, but at the same time it drives the process of fibrosis in the deeper layers of the gut (Van Assche *et al.* 2004). Nonetheless, we could not find an association between the IVS-128A>T polymorphism and stricturing disease, indicating that additional risk factors influence this particular phenotype.

Interestingly, though, the IVS-128T allele in *MTF1* is associated with CD of the ileum, which has a higher likelihood of stricturing than does colonic CD.

We used microarray screenings in unaffected tissues for human genetic studies. A similar study was performed by Lawrance and colleagues, with the exception that they used moderately inflamed resected colonic tissue (Lawrance *et al.* 2001). We used non-inflamed tissue to target basal differences in gene expression due to genetic variation, and not due to inflammation-related events. Significant evidence was found for the genetic transmission of variation in gene expression (Lo *et al.* 2003; Pastinen *et al.* 2004). The variation in expression level is highest among unrelated individuals, and smallest between monozygotic twins, indicating that germ-line differences contribute to variation in gene expression (Cheung *et al.* 2003; Correa and Cheung 2004).

In summary, screening unaffected colon biopsies by microarrays proved to be useful in the identification of new candidate genes for CD. A cluster of MT genes located at *IBD1* is consistently down-regulated in CD patients with colonic involvement. Subsequently, we identified a new disease-modifying gene, *MTF1*, that is associated with ileal disease, and together with *CARD15* serves as a good predictor of disease location in CD patients.

AKNOWLEDGEMENTS

We acknowledge the contribution of the VIB Genetic Service Facility (http://www.vibgeneticservicefacility.be) to the genetic analyses, and the Protein expression and Purification core facility of VIB for providing recombinant TNF.

REFERENCES

Ahmad, T., A. Armuzzi, M. Bunce*, et al.* (2002). "The molecular classification of the clinical manifestations of Crohn's disease." Gastroenterology. **122**(4): 854-66.

Boudreau, F., E. H. Rings, H. M. van Wering*, et al.* (2002). "Hepatocyte nuclear factor-1 alpha, GATA-4, and caudal related homeodomain protein Cdx2 interact functionally to modulate intestinal gene transcription. Implication for the developmental regulation of the sucrase-isomaltase gene." J Biol Chem. **277**(35): 31909-17.

Brant, S. R. and Y. Y. Shugart (2004a). "Inflammatory bowel disease gene hunting by linkage analysis: rationale, methodology, and present status of the field." Inflamm Bowel Dis **10**(3): 300-11.

Brant, S. R. and Y. Y. Shugart (2004b). "Inflammatory bowel disease gene hunting by linkage analysis: rationale, methodology, and present status of the field." Inflamm Bowel Dis.

10(3): 300-11.

Bruwer, M., K. W. Schmid, K. A. Metz*, et al.* (2001). "Increased expression of metallothionein in inflammatory bowel disease." Inflamm Res. **50**(6): 289-93.

Cheung, V. G., L. K. Conlin, T. M. Weber*, et al.* (2003). "Natural variation in human gene expression assessed in lymphoblastoid cells." Nat Genet **33**(3): 422-5..

Cheung, V. G. and R. S. Spielman (2002). "The genetics of variation in gene expression." Nat Genet **32**(Suppl): 522-5.

Clarkson, J. P., M. E. Elmes, B. Jasani*, et al.* (1985). "Histological demonstration of immunoreactive zinc metallothionein in liver and ileum of rat and man." Histochem J. **17**(3): 343-52.

Correa, C. R. and V. G. Cheung (2004). "Genetic variation in radiation-induced expression phenotypes." Am J Hum Genet **75**(5): 885-90.

Cuthbert, A. P., S. A. Fisher, M. M. Mirza*, et al.* (2002). "The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease." Gastroenterology. **122**(4): 867-74.

Darfeuille-Michaud, A., J. Boudeau, P. Bulois*, et al.* (2004). "High prevalence of adherent-invasive Escherichia coli associated with ileal mucosa in Crohn's disease." Gastroenterology. **127**(2): 412-21.

Dooley, T. P., E. V. Curto, S. P. Reddy*, et al.* (2004). "Regulation of gene expression in inflammatory bowel disease and correlation with IBD drugs: screening by DNA microarrays." Inflamm Bowel Dis **10**(1): 1-14.

Eckmann, L. and M. Karin (2005). "NOD2 and Crohn's disease: loss or gain of function?" Immunity. **22**(6): 661-7.

Elmes, M. E., J. P. Clarkson and B. Jasani (1986). "The immunocytochemical demonstration of metallothionein in human liver and small intestine." Acta Pharmacol Toxicol (Copenh). **59**(Suppl 7): 510-3.

Gasche, C., B. Z. Alizadeh and A. S. Pena (2003). "Genotype-phenotype correlations: how many disorders constitute inflammatory bowel disease?" Eur J Gastroenterol Hepatol. **15**(6): 599-606.

Gasche, C., J. Scholmerich, J. Brynskov*, et al.* (2000). "A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998." Inflamm Bowel Dis. **6**(1): 8-15.

Halfvarson, J., L. Bodin, C. Tysk*, et al.* (2003). "Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics." Gastroenterology **124**(7): 1767-73.

Hampe, J., H. Frenzel, M. M. Mirza*, et al.* (2002). "Evidence for a NOD2-independent

susceptibility locus for inflammatory bowel disease on chromosome 16p." Proc Natl Acad Sci U S A. **99**(1): 321-6.

Haq, F., M. Mahoney and J. Koropatnick (2003). "Signaling events for metallothionein induction." Mutat Res. **533**(1-2): 211-26.

Haroon, Z. A., K. Amin, P. Lichtlen*, et al.* (2004). "Loss of metal transcription factor-1 suppresses tumor growth through enhanced matrix deposition." Faseb J. **18**(11): 1176-84.

Hugot, J. P., C. Alberti, D. Berrebi*, et al.* (2003). "Crohn's disease: the cold chain hypothesis." Lancet. **362**(9400): 2012-5.

Hugot, J. P., M. Chamaillard, H. Zouali*, et al.* (2001). "Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease." Nature **411**(6837): 599-603.

Ioachim, E., M. Michael, C. Katsanos*, et al.* (2003). "The immunohistochemical expression of metallothionein in inflammatory bowel disease. Correlation with HLA-DR antigen expression, lymphocyte subpopulations and proliferation-associated indices." Histol Histopathol. **18**(1): 75-82.

Itoh, N., H. Shibayama, M. Kanekiyo*, et al.* (2005). "Reduced bactericidal activity and nitric oxide production in metallothionein-deficient macrophages in response to lipopolysaccharide stimulation." Toxicology. **216**(2-3): 188-96.

Karin, M., R. L. Eddy, W. M. Henry*, et al.* (1984). "Human metallothionein genes are clustered on chromosome 16." Proc Natl Acad Sci U S A. **81**(17): 5494-8.

Kimura, T., N. Itoh, M. Takehara*, et al.* (2002). "MRE-binding transcription factor-1 is activated during endotoxemia: a central role for metallothionein." Toxicol Lett. **129**(1-2): 77-84.

Kleinjan, D. A. and V. van Heyningen (2005). "Long-range control of gene expression: emerging mechanisms and disruption in disease." Am J Hum Genet **76**(1): 8-32.

Kruidenier, L., I. Kuiper, W. Van Duijn*, et al.* (2003). "Imbalanced secondary mucosal antioxidant response in inflammatory bowel disease." J Pathol. **201**(1): 17-27.

Laukens, D., H. Peeters, D. Marichal*, et al.* (2005). "CARD15 gene polymorphisms in patients with spondyloarthropathies identify a specific phenotype previously related to Crohn's disease." Ann Rheum Dis. **64**(6): 930-5.

Laukens, D., H. Peeters, B. Vander Cruyssen*, et al.* (2006). "Altered gut transcriptome in spondyloarthropathy." Ann Rheum Dis **13**.

Lawrance, I. C., C. Fiocchi and S. Chakravarti (2001). "Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes." Hum Mol Genet. **10**(5): 445-56.

Lesage, S., H. Zouali, J. P. Cezard*, et al.* (2002). "CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease." Am J Hum Genet. **70**(4): 845-57.

Lo, H. S., Z. Wang, Y. Hu*, et al.* (2003). "Allelic variation in gene expression is common in the human genome." Genome Res **13**(8): 1855-62.

Marks, D. J., M. W. Harbord, R. MacAllister*, et al.* (2006). "Defective acute inflammation in Crohn's disease: a clinical investigation." Lancet **367**(9511): 668-78.

Mulder, T. P., H. W. Verspaget, A. R. Janssens*, et al.* (1991). "Decrease in two intestinal copper/zinc containing proteins with antioxidant function in inflammatory bowel disease." Gut. **32**(10): 1146-50.

Ogura, Y., D. K. Bonen, N. Inohara*, et al.* (2001). "A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease." Nature **411**(6837): 603-6.

Pastinen, T., R. Sladek, S. Gurd*, et al.* (2004). "A survey of genetic and epigenetic variation affecting human gene expression." Physiol Genomics **16**(2): 184-93.

Peeters, M., H. Nevens, F. Baert*, et al.* (1996). "Familial aggregation in Crohn's disease: increased age-adjusted risk and concordance in clinical characteristics." Gastroenterology. **111**(3): 597-603.

Puskas, L. G., A. Zvara, L. Hackler, Jr.*, et al.* (2002). "RNA amplification results in reproducible microarray data with slight ratio bias." Biotechniques **32**(6): 1330-4, 1336, 1338, 1340.

Samson, S. L. and L. Gedamu (1998). "Molecular analyses of metallothionein gene regulation." Prog Nucleic Acid Res Mol Biol. **59**: 257-88.

Shaw, S. H., J. Hampe, R. White*, et al.* (2003). "Stratification by CARD15 variant genotype in a genome-wide search for inflammatory bowel disease susceptibility loci." Hum Genet. **113**(6): 514-21.

Smith, E., P. A. Drew, Z. Q. Tian*, et al.* (2005). "Metallothionien 3 expression is frequently down-regulated in oesophageal squamous cell carcinoma by DNA methylation." Mol Cancer. **4**: 42.

Sturniolo, G. C., C. Mestriner, P. E. Lecis*, et al.* (1998). "Altered plasma and mucosal concentrations of trace elements and antioxidants in active ulcerative colitis." Scand J Gastroenterol. **33**(6): 644-9.

Van Assche, G., K. Geboes and P. Rutgeerts (2004). "Medical therapy for Crohn's disease strictures." Inflamm Bowel Dis. **10**(1): 55-60.

van Heel, D. A., B. M. Dechairo, G. Dawson*, et al.* (2003). "The IBD6 Crohn's disease locus demonstrates complex interactions with CARD15 and IBD5 disease-associated variants." Hum Mol Genet. **12**(20): 2569-75.

Vandesompele, J., K. De Preter, F. Pattyn*, et al.* (2002). "Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes." Genome Biol. **3**(7): RESEARCH0034.

Vermeire, S., P. Rutgeerts, K. Van Steen*, et al.* (2004a). "Genome wide scan in a Flemish inflammatory bowel disease population: support for the IBD4 locus, population heterogeneity, and epistasis." Gut **53**(7): 980-6.

Vermeire, S., G. Van Assche and P. Rutgeerts (2004b). "C-reactive protein as a marker for inflammatory bowel disease." Inflamm Bowel Dis. **10**(5): 661-5.

Weckx, S., P. De Rijk, C. Van Broeckhoven*, et al.* (2005a). "SNPbox: a modular software package for large-scale primer design." Bioinformatics. **21**(3): 385-7.

Weckx, S., J. Del-Favero, R. Rademakers*, et al.* (2005b). "novoSNP, a novel computational tool for sequence variation discovery." Genome Res. **15**(3): 436-42.

Zangger, K., G. Oz, E. Haslinger*, et al.* (2001). "Nitric oxide selectively releases metals from the amino-terminal domain of metallothioneins: potential role at inflammatory sites." Faseb J. **15**(7): 1303-5.

Zhang, B., O. Georgiev, M. Hagmann*, et al.* (2003). "Activity of metal-responsive transcription factor 1 by toxic heavy metals and H2O2 in vitro is modulated by metallothionein." Mol Cell Biol. **23**(23): 8471-85.

Chapter 4

HUMAN METALLOTHIONEIN EXPRESSION UNDER NORMAL AND PATHOLOGICAL CONDITIONS: MECHANISMS OF GENE-REGULATION

Debby Laukens[1], Anouk Waeytens[2], Pieter De Bleser[1], Claude Cuvelier[2], Martine De Vos[3] and Pieter Rottiers[1]

[1]Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium
[2]Department of Pathology, Ghent University Hospital, Gent, Belgium
[3]Department of Gastroenterology, Ghent University Hospital, Gent, Belgium

ABSTRACT

Metallothioneins (MT) are ubiquitous metal-binding proteins that are highly conserved throughout evolution. Although the exact physiological function is not completely understood, it is clear that they are involved in a variety of processes including metal detoxification, free radical scavenging, metal homeostasis and cell proliferation. The human MT family consists of at least 18 different isoforms, containing pseudogenes as well as functional proteins. They can be induced by a wide variety of substances, e.g. metals, cytokines and hormones. In addition, different cell types express discrete MT isoforms, reflecting the specifically adapted functions of MT isoforms, and hence a divergence in their regulation. Aberrant expression of MT has been described in a number of apparently diverse diseases, including inflammatory bowel disease, cancer, Alzheimer's disease, amyotrophic lateral sclerosis and Menkes disease. Therefore, a thorough understanding of the regulation of MT expression is important. To date, the regulation of transcription of these genes has primarily been studied in mice. Unfortunately, the situation in mice is somehow less complicated, since only 4 isoforms are expressed. Nevertheless, the high homology between mouse and human MTs allows us to evaluate regulatory regions in their respective promoters. Here, we review the aberrant expression of human MT in disease, and the mechanisms that regulate MT expression.

INTRODUCTION

Metallothioneins (MT) are a family of small, highly conserved proteins with the specific capacity to bind metal ions. The MT protein was first purified from the equine renal cortex in 1960 (Kagi and Valee 1960). Since, they were described in a wide variety of species, including vertebrates, invertebrates, plants, fungi and some prokaryotes. A great deal of sequence and structural homology exists between MT proteins in different species, underlining its important biological role. Mammalian MT proteins typically consist of 61 to 68 amino acids, with a high content of polar, highly catalytic cysteine residues. These cysteins are strictly conserved and arranged in motifs that form the framework of two distinct metal-binding domains, linked by a short peptide (Figure 1). Initially, the biological function of MT was centralized to their unique metal-binding capacity. In normal conditions, excessive concentrations of essential and nonessential metal ions like cadmium, mercury, and lead can be toxic. Most organisms use a redundant array of cellular mechanisms to limit toxicity of metal ions (Dameron and Harrison 1998), one of which is sequestration by MT. They thus provide homeostasis of metal ions and protect cells from acute heavy metal toxicity. Through participating in zinc metabolism, they regulate the activity of fast exchanging metalloproteins, such as NFκB (Sakurai *et al.* 1999; Kim *et al.* 2003) and the tumor suppressor gene p53 (Ostrakhovitch *et al.* 2006). It is now clear that the function of MT is not restricted to this metal-binding activity. They are rapidly induced during specific phases of the cell cycle and by diverse stimuli, such as metals, hormones and cytokines, thus they participate in cell cycle and diverse protective functions. Metallothioneins are capable of scavenging free radicals and thus play a role in protection of tissues against various forms of oxidative injury, including radiation, lipid peroxidation, oxidative stress conditions of hyperoxia. In addition, extracellular MT has been shown. Elevated serum and urine MT was found in cadmium-exposed individuals (Falck *et al.* 1983). Moreover, MT in the extracellular environment may act as a "danger-signal" promoting movement of leukocytes to the site of inflammation (Yin *et al.* 2005). A significant role for a strictly controlled regulation of MT in both growth of cells and in their response towards several stimuli is obvious. It is therefore plausible that defects in the one of these restricted mechanisms can lead to pathological situations, like for example tumor growth. Good knowledge of their regulation is thus crucial in understanding their role in pathogenesis.

THE HUMAN METALLOTHIONEIN FAMILY

The classification of metallothioneins has been revised in 1999, and is based on evolutionary

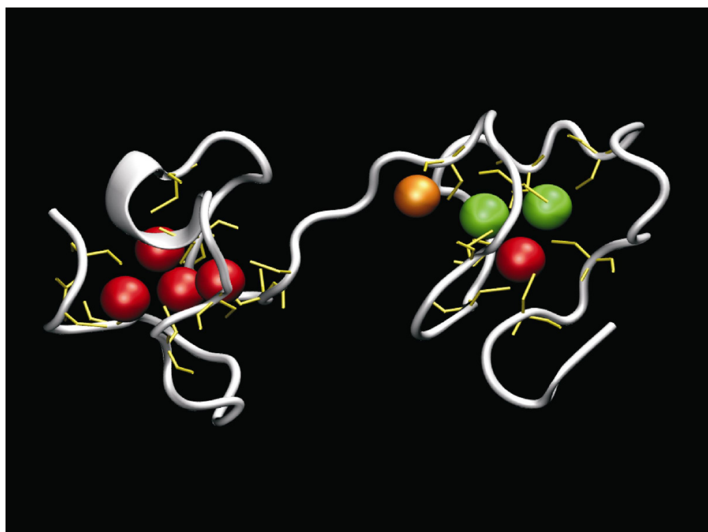


Figure 1 Homology model of human metallothionein 1A. Cysteine residues are shown as yellow sticks, metal ions are shown in red (Cd^{2+}), green (Zn^{2+}) and orange (Na^{+}). The model was constructed using the Swiss Model web server (Schwede et al. 2003) with the crystal structure of the rat metallothionein 1A protein (PDB ID 4MT2) as template structure.

data (Binz 1999). Rodents have four MT isoforms, MT1 to MT4, while all primates examined so far contain multiple copies of the MT1 isoform. The situation is most complex in humans: a total of 18 MT isoforms and 5 MT-like genes have been cloned so far, many of which only differ in distinct amino acids (Table 1). At least five of the isoforms are non-processed pseudogenes. Seventeen out of 18 isoforms cluster together on chromosome 16q13 (Figure 2) (Karin *et al.* 1984). Apparently, chromosome 16 is one of the most enriched chromosomes for intrachromosomal duplications as compared to the human genomic average (Martin *et al.* 2004), and includes for example the cadherin gene cluster. It is not clear whether some of the MT pseudogenes are functional: MT1J and MT1M contain a promoter, are transcribed, but contain a premature stop codon.

The various MT isoforms differ mainly in their expression pattern: MT3 an MT4 are constitutively expressed in specific cell types, while MT1 and MT2 are highly inducible and ubiquitously expressed (Table 1).

By comparing the coding sequences of rodent and primate MT genes, it was shown that human MT1 isoforms show less divergence from human MT2 than from mouse MT1, suggesting that these two isoforms arose after the emergence of primates (Schmidt *et al.* 1985). Similarly, MT3 and MT4 probably diverged prior to the primate/rodent divergence (Figure 3). In addition, MT1 and MT2 isoforms are clearly separated from the MT3 and MT4 clusters, probably reflecting their unique functions in the organism. Moreover, this means that MT1 and MT2 proteins might have a very similar function in mice and in humans.

Table 1 Annotated human metallothionein genes and their characteristics

Symbol	Chromosomal location	Aliases	Protein length	Refseq status	Expression
MT:					
MT4	16q13		62	provisional	squamous epithelium
MT3	16q13	GIF	68	validated	brain, kidney, reproductive system
MT2A	16q13	MT2	61	provisional	ubiquitous
MT1A	16q13	MT1S	61	validated	ubiquitous
MT1B	16q13	MT1Q	61	provisional	ubiquitous
MT1E	16q13		61	validated	ubiquitous
MT1F	16q13		61	provisional	ubiquitous
MT1G	16q13		61	provisional	ubiquitous
MT1H	16q13	MT-0	61	provisional	ubiquitous
MT1J	16q13	MT1NP	40	provisional	?
MT1M	16q13	MT1K	61	provisional	ubiquitous
MTM	16q13		49	provisional	?
MT1X	16q13		61	provisional	ubiquitous
MT pseudogenes:					
MT1I	16q13?		-	withdrawn	-
MT1L	16q13	MT1R	-	provisional	-
MT1CP	16q13		-		-
MT1DP	16q13		-		-
MT2P1	4q13		-	provisional	-
MT-like:					
MTL1	1p		?		?
MTL2	1p22		?		?
MTL3	18		?		?
MTL4	20		?		?
MTL5	11q13	TESMIN	178	reviewed	testis

METALLOTHIONEIN IN DISEASE

Metallothionein expression has been studied in various pathological conditions. Unfortunately, contradictory results complicate the interpretation of the involvement of these proteins in diseases. Expression data obtained by immunohistochemistry, semi-quantitative PCR and more

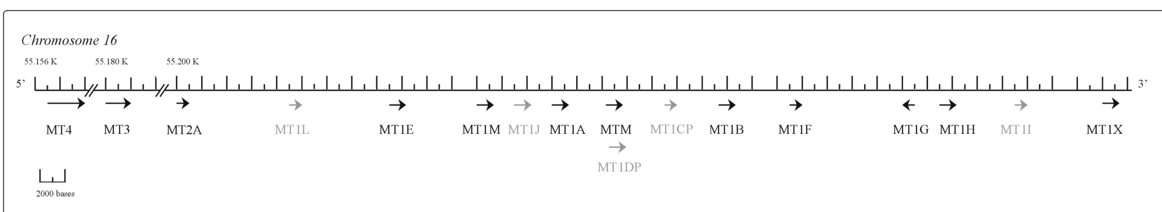


Figure 2 Genomic organization of the human metallothionein family on chromosome 16. Pseudogenes are represented in grey.

recent techniques such as quantitative PCR (qPCR) are not always comparable. The early studies relied on immunohistochemistry using antibodies, which cannot distinguish between MT isoforms, because they share the antigenic epitope. Furthermore, in normal tissues, MT expression is usually undetectable by immunohistochemistry, except in myoepithelial (van den Oord and De Ley 1994), renal (Mitropoulos *et al.* 2005), intestinal (Laukens *et al.* 2006) and thyroid epithelial cells, pancreas (Tomita and Matsubara 2000) and fetal liver (Fuller *et al.* 1990). On the other hand, strong staining was found in pathological tissues, especially in many tumors (Theocharis *et al.* 2004). Here, we summarize data of deviant MT expression in several diseases.

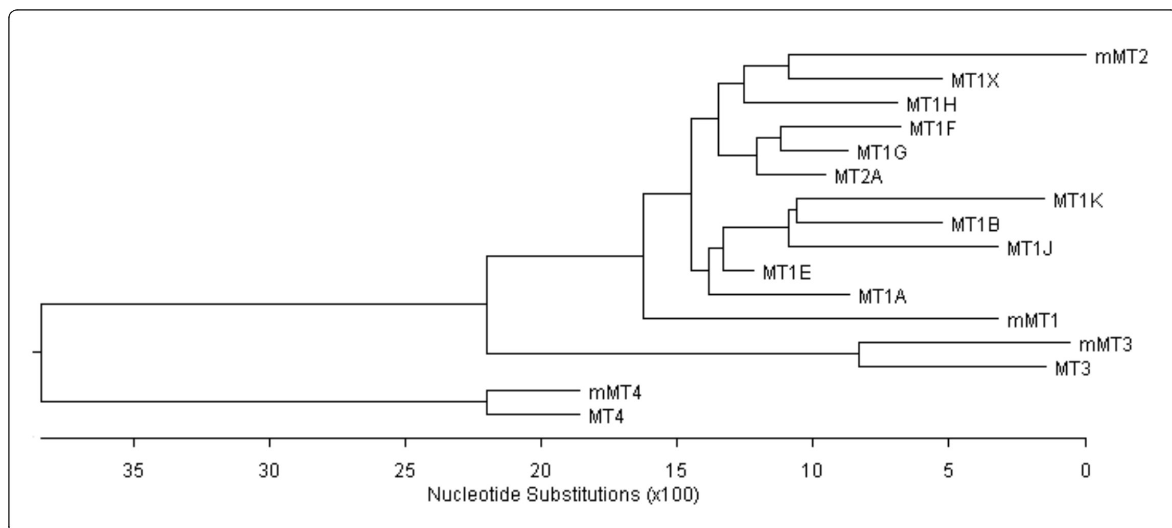


Figure 3 Phylogenetic tree of human and mouse metallothionein protein isoforms. Metallothionein 3 and 4 probably arose prior to the divergence of primates and rodents, while MT1 and MT2 genes arose later. Mouse MT genes are designated as mMT1, mMT2, mMT3 and mMT4.

Metallothionein and inflammatory bowel disease

Inflammatory bowel disease (IBD), comprises both Crohn's disease (CD) and ulcerative colitis (UC). They are chronic inflammatory diseases of the intestine. In CD, the complete bowel wall is affected, while in UC, only the superficial layers become inflamed. In the normal intestine, we found that MT RNA and protein expression is highest in the ileum as compared to other compartments of the gut (Laukens *et al.* 2006). Metallothionein is localized in the enterocytes, predominantly at the base of crypts, probably due to the high proliferative state of these cells. The immunopositivity decreases towards the top of the villi. Focal staining is often observed in the intestine, as a result of the uniform staining in distinct single crypts. It was suggested that this might be the consequence of somatic mutations in stem cells, leading to strong clonal expression

of MT in the entire crypt (Jasani *et al.* 1998). Aberrant protein expression was found in tissue samples originating from IBD patients. Down-regulation as well as up-regulation has been reported. Two papers reported an up-regulation (Bruwer *et al.* 2001; Dooley *et al.* 2004), while we and others found a down-regulation of MT in IBD (Clarkson *et al.* 1985; Elmes *et al.* 1986; Ioachim *et al.* 2003; Kruidenier *et al.* 2003; Laukens *et al.* 2006). Moreover, MT quantification in IBD based on radioimmunoassay (Mulder *et al.* 1991), microarray (Lawrance *et al.* 2001) and silver-saturation assay (Sturniolo *et al.* 1998) also reported a down-regulation. Using qPCR, we were able to show a decreased expression of a large proportion of MT1 isoforms and MT2 at basal level in colon biopsies, ileum biopsies and whole blood samples of CD patients that have colonic involvement (Laukens *et al.* 2006). This lowered mRNA expression level correlated with protein expression, even though biopsies were not acquired at the same time. These results suggested that the deficient MT expression in CD patients with colonic disease is genetically determined, and not the result of an early inflammation event. Interestingly, the MT gene cluster is located in *IBD1*, a locus that was significantly associated to CD in many genome scans (Mathew and Lewis 2004). In murine IBD models, however, MT does not appear to influence the development or progression of intestinal pathology in the DSS (Oz *et al.* 2005).

Zinc is often decreased in IBD patients. This element plays an important role in the prevention of free radical formation and in protection of biological structures from damage (Stefanidou *et al.* 2006). It was shown that dietary zinc causes an MT increase in all gut regions in rats (Tran *et al.* 1999; Szczurek *et al.* 2001). On the other hand, zinc supplementation in IBD patients did not change concentrations of MT in plasma and erythrocytes (Mulder *et al.* 1994), and the MT concentration in both inflamed and non-inflamed intestinal mucosa was only slightly higher. Histological inflammation scores of intestinal biopsies, plasma albumin levels, and the disease activity index of the patients did not change during the study. However, in this study, only inactive to moderately active patients were included.

A decreased MT level in CD patients could indicate a hampered maintenance of free radicals. Indeed, human monocytic cells that were transfected with an MT-antisense vector, produce more hydrogen peroxide than control THP-1 cells in the absence of a stimulus (Leibbrandt *et al.* 1994). The exact role of MT in CD pathogenesis has to be established in future studies.

Metallothionein and cancer

Immunocytochemically detectable MT overexpression was described in a variety of human tumors, and has been extensively reviewed recently (Theocharis *et al.* 2004). As such, we will not focus on this issue here.

The expression of MT is not universal to all human tumors, but may depend on their differentiation status and proliferative index. Metallothionein overexpression is associated with resistance to anticancer drugs and is combined with a poor prognosis. However, its use as a marker of tumour differentiation, cell proliferation and prognosis predictor remains unclear. On the other hand, gastric carcinomas and colorectal adenomas are apparently accompanied by a decreased expression of MT, however, those with a relatively high level seem to have an increased malignant potential (Janssen *et al.* 2000).

Cell-type specific differential regulation of human MT genes was found in different cancer cell lines, correlating with DNA methylation and chromatin structure (see later) (Jahroudi *et al.* 1990) . Tumor cell-lines arising from paraxial mesoderm and endoderm have MT2A and MT1E genes in fully inducible form and the MT1F in the refractory state (Schmidt *et al.* 1985). On the other hand, tumours originating from ectoderm, intermediate and lateral mesoderm exhibit MT2A and MT1F genes in inducible form and the MT1E gene in a refractory form.

Metallothionein, Alzheimer's disease and amyotrophic lateral sclerosis

Metallothionein 3 was first cloned as growth inhibitory factor (GIF), which showed a decreased RNA expression in Alzheimer's disease (AD). Metallothionein 3 suppresses the neurotrophic activity present in the normal human brain (Tsuji *et al.* 1992b). The down-regulation of MT3 in AD has been confirmed by two recent studies. In the first study, MT3 expression was determined in a large number of AD cases by qPCR as well as by immunohistochemistry and Western blotting (Yu *et al.* 2001). In the second study, DNA microarrays were used to compare RNA levels from control and AD hippocampal regions and found, amongst others, MT3 down-regulation (Colangelo *et al.* 2002). However, Erickson and colleagues (Erickson *et al.* 1994) disputed that neuronal changes in AD are related to a decrease in MT3, since they could not find a significant down-regulation in neither RNA nor protein expression in their AD population. There is more consensus on the overexpression of MT1 and MT2 in the astrocytes from AD as well as other neurological disorders (Duguid *et al.* 1989; Adlard *et al.* 1998; Zambenedetti *et al.* 1998; Chung *et al.* 2004). In the brain, astrocytes are the main source of MT1 and MT2, although other cell types, such as choroid plexus epithelia, endothelium and meningeal cells may also express these isoforms (Penkowa 2006). In neurodegenerative diseases such as AD, astrocytes become abundant and activated in the affected areas. While in other organs, the main function of MT is related to zinc metabolism and protection against heavy metal and/or oxidative damage, the key role of MTs in the brain seems to be a protection in the cellular response to neuronal injury. It was suggested that the specific increase in MTs was associated with the initial

stages of the disease process (Adlard *et al.* 1998). The precise mechanisms downstream of MT have not been fully established, but convincing data showed that they are essential in dealing with neuropathology and for brain recovery in AD as well as other brain pathologies. MTs might even be used as therapeutic and/or preventive drugs for a range of brain disorders (Penkowa 2006).

Amyotrophic lateral sclerosis (ALS) is a progressive, invariably fatal neurological disease due to degeneration of the nerve cells responsible for controlling voluntary muscles. In ALS, motor neurons in the brain stem, spinal cord and motor cortex degenerate or die, ceasing to send messages to muscles. Consequently, the muscles gradually weaken, waste away, and twitch. Fifteen to 20% of cases of familial amyotrophic lateral sclerosis are associated with mutations in the superoxide dismutase 1 gene (*SOD1*). Elevated levels of MT have been found in spinal cord, kidney and liver of patients with ALS (Sillevis Smitt *et al.* 1992a; Sillevis Smitt *et al.* 1992b), but not in serum (Sillevis Smitt *et al.* 1994). A detailed study on different MT isoform expression revealed no evidence for either the induction of a specific MT repertoire, or for the inability of glia to express any MT gene (Blaauwgeers *et al.* 1996). Probably, the enhanced expression of MT in ALS reflects an early protective function. This was also concluded from a study on MT expression in mice carrying the *SOD1* mutation (Gong and Elliott 2000). These mice were backcrossed with MT-knockout mice. The offspring reached the onset of clinical signs significantly earlier in response to the reduction of protein expression. These results indicated that the copper-mediated free radical generation derived from mutant *SOD1* might be related to the degeneration of motor neurons in ALS and that MT might play a protective role against the expression of the disease (Nagano *et al.* 2001).

Recently, MT3 was screened for mutations in 20 patients with ALS, but no functionally relevant polymorphism could be associated with the disease (Morahan *et al.* 2005).

Metallothionein, Menkes disease and Wilson disease

Menkes disease is an X-linked, recessive disorder of the copper metabolism that occurs in less than 1 in 200,000 live births. The condition is characterized by early retardation in growth, peculiar hair, focal cerebral and cerebellar degeneration, skeletal abnormalities, and patient mortality in early childhood (Bankier 1995). Three independent research groups cloned the Menkes gene, a copper transporting ATPase, *ATP7A*, to the long arm of the X chromosome (Chelly *et al.* 1993; Mercer *et al.* 1993; Vulpe *et al.* 1993). A spectrum of mutations adversely affecting protein expression have been observed in severely affected Menkes patients (Kodama and Murata 1999). The genetic defect in Menkes syndrome leads to a progressive copper

deficiency and copper-dependent enzymes fail in most tissues.

Using cultured fibroblasts, it was shown that low extracellular copper concentrations induce synthesis of MT in Menkes' cells but not in normal cells (Hamer 1987). Therefore, in the early studies on tissue abnormalities in Menkes disease, MT were believed to be involved in pathogenesis (Garnica *et al.* 1978; Schmidt *et al.* 1984). Now, it is clear that a defect in efflux and consequently intracellular accumulation of copper, due to mutations in *ATP7A*, is responsible for this up-regulation in Menkes' cells. As a result of accumulation of copper in the cell MTs are up-regulated, but they are as such not causative for pathogenesis in Menkes disease. In affected cells, copper accumulates bound to MT in the cytosol, while its transport to the organelles, as well as copper efflux, is disturbed. The low activity of metalloenzymes is believed to contribute significantly to the pathogenesis of this condition.

Wilson disease is an autosomal recessive disorder characterized by dramatic build-up of intracellular hepatic copper with subsequent hepatic and neurologic abnormalities. Copper toxicity occurs when the liver is overloaded and non-ceruloplasmin-bound copper is released into the bloodstream, from where it can diffuse into the brain. Today, the treatment of Wilson disease is no longer aimed at 'decoppering', the removal of accumulated copper, but at the normalization of the free copper concentration in blood, to reverse the copper poisoning. Therefore, new therapy is aimed at administration of zinc to these patients, because this increases MT expression and sequesters the excess of copper in the blood (Hoogenraad 2006).

HUMAN METALOTHIONEIN GENE REGULATION

Because MT expression is involved in a number of pathological conditions, the transcriptional control of MT has become a major topic. Metallothioneins are induced by a wide variety of physiological and chemical agents like cytokines, metals, hormones, and stress in general (Borghesi and Lynes 1996). Furthermore, they are transiently induced after tissue injury caused by e.g. inflammation or irradiation (Manuel *et al.* 1992). It is generally accepted that *MT1* and *MT2* genes are inducible, while *MT3* and *MT4* are constitutively expressed. Metallothionein 1 isoforms have a restricted and transient role, perhaps in stress or during infection and the cell cycle. On the other hand, MT2A is ubiquitously expressed, and plays a general role in cellular physiology, possibly in zinc metabolism. Metallothionein 3 is predominantly expressed in the central nervous system (Tsuji *et al.* 1992a; Blaauwgeers *et al.* 1996), but also in the kidney (Hoey *et al.* 1997), prostate (Dutta *et al.* 2002), retina (Tate *et al.* 2002), salivary glands (Irie *et al.* 2004) and reproductive system (Moffatt and Seguin 1998). The mouse MT4 is expressed in

stratified squamous epithelium of the tongue (Quaife *et al.* 1994), but there are so far no data on human MT4 expression.

Most of the functional studies on MT transcription were performed in mice. However, in mice, MT1 and MT2 isoforms are co-ordinately regulated (Searle *et al.* 1984), while the many human MT isoforms are regulated in a cell-type specific manner (Schmidt and Hamer 1986; Varshney *et al.* 1986; Laukens *et al.* 2006). This probably resulted from the adaptation of the MT isoforms to more specific functions throughout evolution. Phylogenetic tree analysis of the promoter region of mouse and human MT isoforms shows that MT2, MT3 and MT4 promoters are highly homologue between mouse and human, which might reflect their more strictly regulated expression pattern. Human MT1 isoforms are dispersed into two groups. Interestingly, these two clusters correlate to the physical location of the genes on chromosome 16 (Figure 2). On the other hand, the mouse MT1 promoter region is relatively unrelated to the human MT1 group (Figure 4). It is thus plausible that regulatory regions in the mouse and human promoter have adapted differently. Therefore, care must be taken when evaluating MT expression and maybe even function in the mouse, for solving questions related to human circumstances. Nevertheless, if common transcription factor binding sites are found in human and mouse promoters, we can extrapolate the experimental data on mouse MT1 regulation to the human situation.

We performed a comparative study on the transcriptional control of MT in mice and humans, using *in silico* data of promoter transcription factor binding sites. Based on these common regulatory sites, we review the MT regulatory mechanisms.

In silico analysis of human and mouse MT1 promoters

Although marked sequence conservation exists, the RNA level for each MT isoform is unique. This is possibly due to inherent differences in promoter regulatory sequences. Promoters are organized with a variety of elements that contribute to promoter function. The elements found in any promoter differ in number, location and orientation.

A difference distance matrix approach (De Bleser *et al.* 2006) was used to identify a set of transcription binding sites that are specifically highly present in following MT isoforms: MT1A, MT1B, MT1E, MT1F, MT1G, MT1E, MT1J, MT1M and MT1X (Table 2). Many of these binding sites and their transcription factors have been described in MT regulation, however, we found a number of unexplored regulatory regions. We next searched for these transcription factor binding sites in the mouse MT1 promoter, and for well-known mouse binding sites (USF, MLTF) in the promoters of human MTs. Below, we describe the involvement of these transcriptional regulators in more detail.

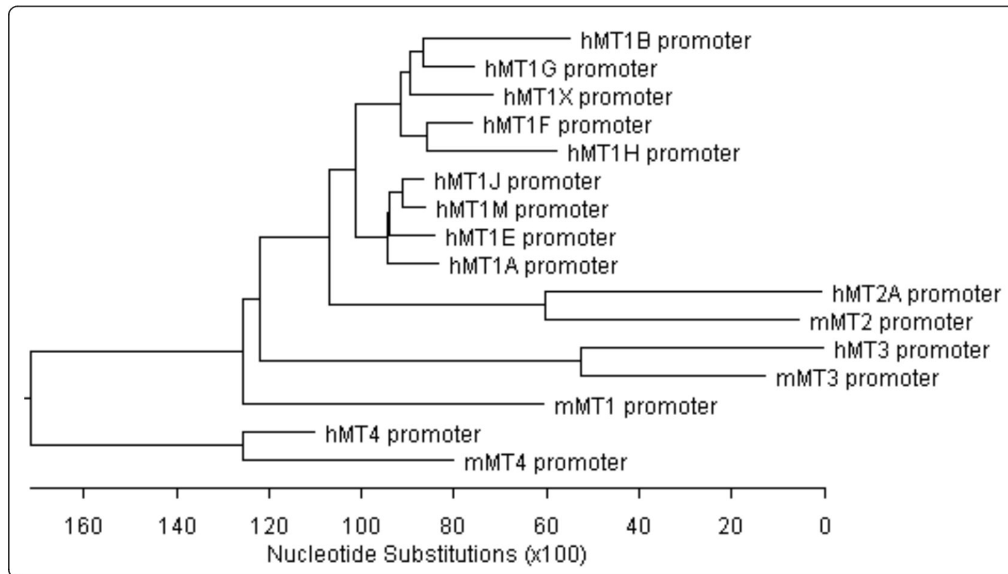


Figure 4 Phylogenetic tree analysis of the promoter regions of human and mouse MT isoforms. Promoter regions are defined from nucleotides -800 to -1 relative to the transcriptional start site.

Table 2 Transcription factor binding sites in 800 bp of the promoter of human and mouse metallothionein 1 isoforms

	Transcription factor	MT1A	MT1B	MT1E	MT1F	MT1G	MT1H	MT1J	MT1M	MT1X	mMT1
basal	TBP	+	+	-	-	+	-	-	-	+	+
	TFII-I	2	2	1	3	1	2	2	4	0	1
	Sp1	5	1	1	8	5	2	2	4	3	3
	AP2	0	1	1	1	1	3	1	0	0	0
	USF/Nrf2	0	0	0	0	1	0	0	0	0	1
induction	ChCh	9	3	6	8	6	5	8	9	5	3
	HELIOS	2	8	2	6	6	7	4	3	9	3
	E2F	3	0	4	4	3	2	6	5	3	3
	Spz1	4	4	5	4	2	5	3	6	1	2
	Egr_1	4	0	3	5	0	0	3	8	4	1
	GR	1	1	6	0	0	3	1	2	2	3
	STAT	1	0	4	2	1	1	1	1	2	0
	MTF1	2	4	1	5	2	5	2	2	3	3
	RAR-a	2	5	1	5	3	3	3	2	3	3
repression	ETF	6	2	3	9	6	2	6	6	6	4
	NFI	1	0	1	3	4	2	1	2	2	4

The number of hits as found by Match^{TM} (TRANSFAC PRO 8.4) using a core matrix match of 100% and a matrix match of 85%. Exceptions are IL6RE and MLTF for which consensus sequences were used in combination with Patch^{TM} (TRANSFAC PRO 8.4). The Match procedure introduces false positives; therefore, the number of hits listed is an overestimation of the real hits. The MT isoforms with TBP consensus sequences located close tot the transcription start site are denoted with a plus (+). Those that do not contain a TBP close to the transcription start site are denoted with a minus (-). mMT1= mouse MT1

Basal expression

RNA polymerase II is responsible for transcribing genes coding for messenger RNA (mRNA). The first step in transcription is the binding of the TFIID complex to a region upstream of a sequence called the TATA box. The TFIID complex consists of the TATA-binding protein (TBP) and TBP-associated factors (TAF). The location of the TATA box with respect to the start point is relatively fixed, usually located ~25 bp upstream of the transcriptional start site. Therefore, TATA box consensus sequences located further upstream are probably not functional. Metallothionein 1E, MT1F, MT1H, MT1J and MT1M contain TATA-less promoters. This could partly explain the observation that the MT1G promoter is five times more active than the MT1F promoter in transfection studies (Gedamu *et al.* 1987; Shworak *et al.* 1993). GC boxes, common promoter components involved in basal transcription, are frequently found in MT promoters, often clustered together close to the start site. It has been proposed that the Sp1 factor binds to multiple GC boxes, resulting in an interaction of GC box-bound Sp1 factors with each other to synergistically stimulate transcription. Furthermore, basal level enhancer sequences, binding activator protein-2 (AP2), are found in some human MTs, but not in the mouse MT1. TFII-I binds specifically to initiator elements (Inr), supporting basal transcription. An E box, binding the upstream stimulatory factor (USF) was only found in the mouse MT1 and the human MT1G. It was shown that TFII-I also binds to upstream E box, and that TFII-I and USF interact co-operatively at both Inr and E box sites (Roy *et al.* 1991). The USF sequence overlaps with an antioxidant response element (ARE). These AREs are usually found in genes responsive to free radicals, through interaction of ARE with NF-E2-related factor 2 (Nrf2). However, this sequence is only found in the mouse MT1 and human MT1G, suggesting that they might be more responsive to free radical exposure. Nevertheless, response to oxidative stress can also be mediated by metal responsive elements (MRE) in the promoter (Dalton *et al.* 1994). MREs are recognized by the MRE-binding transcription factor 1 (MTF1), and are classically required for metal induction of MTs (see below). Nevertheless, they also participate in basal transcription. Indeed, the basal expression of MT is highly correlated to the activity of MTF1 (Ghoshal and Jacob 2001). Furthermore, we have recently shown that basal MT levels in whole blood vary considerably between individuals, but they correlate well with MTF1 expression (Laukens *et al.* 2006).

Inducible expression

Metal induction. Metallothionein 1 and MT2 isoforms are highly inducible by many metal ions, including zinc, cadmium, bismuth, mercury, copper, nickel and cobalt. The concentration of

metal ions to induce MTs depends on the type of ion and MT isoform. Metallothioneins are capable of binding most of these elements, however, they do not bind nickel and cobalt. The transcriptional regulation responsible for this metal induced expression, is controlled by MREs, present in the promoter of MT as multiple, non-identical copies. The mouse and human transcription factor that binds to these MRE elements is the MRE-binding transcription factor 1 or MTF1. When this factor was first cloned (Brugnera *et al.* 1994), it was shown that the human MTF1 was more effective than the mouse equivalent. MTF1 is absolutely necessary for both basal and metal inducible MT expression (Heuchel *et al.* 1994). Moreover, although MTF1 is activated by a number of metals, it absolutely requires zinc for its activity. It was hypothesized that inducing non-zinc metals can displace zinc from its storage proteins, resulting in a pool of free zinc available for activation of MTF1 (Jacob *et al.* 1998). These storage proteins could be MT itself, leading to a complex feedback interaction between MTF1 and MTs.

Interestingly, treatment of cells with cadmium increases MT expression, although it does not influence the DNA binding activity of MTF1 to the MRE. This suggests that additional mechanisms play a role in metal induction. For example, cadmium induces oxidative stress, which could activate binding of USF to the E box. Alternatively, as stated before, cadmium might replace the intracellular zinc in storage proteins, which results in more free zinc to activate MTF1.

In addition, posttranslational modification of MTF1 has been shown. Phosphorylation of MTF1 plays a critical role in its activation by zinc and cadmium (Saydam *et al.* 2002). Several phosphorylation sites are present throughout the complete MTF1 protein. This was thought to be mediated through a complex pathway involving protein kinase C, tyrosine kinase, and casein kinase II.

Stress and inflammation mediated induction. Similar to acute phase proteins, MT is induced by inflammation, bacterial infection and stress. Stress in general often results in the synthesis of glucocorticoid hormones, resulting in the suppression of inflammation and an increase in blood sugar levels. These steroid hormones are synthesized by the adrenal gland, and can enter the cell by simple diffusion. Within the cell, it binds to its receptor, the glucocorticoid receptor (GR), which in turn gets activated and translocates to the nucleus. There, it has high affinity for a consensus sequence called the glucocorticoid response element (GRE), and activates transcription from MTs and other GRE containing genes. Glucocorticoid hormones have been known for a long time as inducers of MT (Karin and Herschman 1979; Jacob *et al.* 1999). The human MT2 gene contains one GRE, while the mouse MT1 and MT2 genes contain two tandem

copies ~1kb upstream of the MT2 gene and ~7 kb of the MT1 gene (Kelly *et al.* 1997). In humans, MT2A is significantly more inducible by glucocorticoids as compared to MT1 isoforms (Schmidt and Hamer 1986). It was shown that dexamethasone, a synthetic glucocorticoid agonist, appears to have no significant effect on the expression of MT1F (Varshney *et al.* 1986). This can be explained by the fact that there is no GRE sequence in the promoter of MT1F.

During acute inflammation, such as after tissue damage or during infection, macrophages are recruited and activated to the site of inflammation. They secrete pro-inflammatory cytokines such as IL6, TNF, IL1α and IL1β. These cytokines are able to induce MT (De *et al.* 1990). Fast MT up-regulation after challenge with IL1 is probably mediated through glucocorticoids (Coto *et al.* 1992). Interleukin 6 is one of the most potent inducers of MT (Schroeder and Cousins 1990). Metallothioneins have IL6 response elements (IL6RE) in their promoter. These DNA stretches bind STAT transcription factors. A synergistic effect was demonstrated between IL6 and glucocorticoid in MT induction (Kasutani *et al.* 1998), which is possibly mediated by the close proximity of the GRE and the IL6RE in the MT promoters.

Interestingly, it was reported that MTs inhibit the release of pro-inflammatory cytokines (Kanekiyo *et al.* 2002; Inoue *et al.* 2006). This could be explained by the regulatory role of MT in NFκB activation (Sakurai *et al.* 1999).

Cell cycle. Metallothioneins appear to play a major role during the cell cycle. A tenfold rise of MT synthesis was described in exponentially growing human hepatocytes (Nagel and Vallee 1995; Studer *et al.* 1997), and peaks of MT expression were found in late G1 and G1/S transition in HT29 epithelial cells (Nagel and Vallee 1995). Similarly, in placental tissue, positive immunostaining for MT was found only in trophoblast and proliferating cells (Haerslev *et al.* 1995). We have recently described that MT expression in the intestine was most apparent in the rapidly proliferating cells of the crypts (Laukens *et al.* 2006). Moreover, it was suggested that MT transcription is altered by the differentiation process. Indeed, it was shown that the differentiation of teratocarcinoma cells using retinoic acid is associated with a rise in MT expression. Retinoic acid receptors (RAR) are nuclear receptors related to the steroid and thyroid hormone receptors, a family of proteins that functions as ligand-dependent transcription factors. Retinoic acid is a regulator of differentiation at various stages of vertebrate embryogenesis. In accordance, multiple RAR receptor binding sites are found in human and mouse MT1.

Although MT expression is generally cytosolic, nuclear translocation has been observed at G0/G1 to early S-phase (Cherian and Apostolova 2000). This nuclear and cytoplasmic localization of MT was also observed in several tumours, especially in regions of high

proliferation. Moreover, antisense down-regulation of MT1 in endothelial cells resulted in the cell cycle arrest at the G1 phase (Miyashita and Sato 2005), and cell growth was inhibited in MT1 antisense tumor cells (Takeda *et al.* 1997). Hesketh and colleagues have shown that the nuclear translocation is determined by the 3' untranslated region of MT1 (Hesketh 2004). Recently, they identified a 11 nucleotide sequence in the 3'UTR, containing a CACC repeat, that is necessary for the nuclear translocation of MT1 (Nury *et al.* 2005).

In the human and mouse MT1 promoters, potential binding sites for the E2F transcription factor are present. This transcription factor is a critical determinant of the G1/S-phase transition during the mammalian cell cycle, serving to activate the transcription of a group of genes that encode proteins necessary for DNA replication. In addition, E2F activity appears to be directly regulated by the action of retinoblastoma protein (Rb). Human DP-1 and E2F-1 associate both *in vivo* and *in vitro*, and this interaction leads to enhanced binding to E2F DNA-binding sites (Helin *et al.* 1993). The association of E2F-1 and DP-1 leads to co-operative activation of an E2F-responsive promoter. It was also demonstrated that trans-activation by E2F-1/DP-1 heterodimers is inhibited by RB. Nevertheless, the actual binding of E2F to MT promoters needs to be determined.

The nuclear need for MT at specific stages of the cell cycle might point towards a critical function of MT in regulating metalloproteins, or protection from DNA damage and apoptosis (Meplan *et al.* 1999).

Development. Metallothionein expression is tightly regulated and activated during mammalian embryonic development. During early development of the mouse embryo, expression of MT1 is induced specifically in the endoderm cells of the visceral yolk sac (Andrews *et al.* 2001). It was shown that MTF1 is absolutely essential for up-regulation of MT1 gene expression in visceral endoderm cells and that optimal expression also involves the binding of USF to the promoter. Only the human MT1G promoter contains an E box, therefore, it is not known whether the human USF binding is necessary in development.

Inhibition of transcription

Contrary to the activation of MT expression, their down-regulation by *cis*-acting events has not been extensively studied. However, a reduced expression has been frequently found in many types of cancers and in IBD. Three factors are probably involved in suppression of MT: nuclear factor 1 (NFI), ZBTB11 and EGFR-specific transcription factor (ETF). Overexpression of NFI in human hepatoma cells suppressed both constitutive and metal induced activation of the MT1 promoter (Majumder *et al.* 2001). NFI binds to an MRE-c' sequence in the mouse promoter

(Datta and Jacob 1993). NFI sites are present in the mouse and all human promoters except MT1B. One study reports a 120 kDa zinc finger protein (PZ120) repressing the transcription of the human MT2A gene by binding to its transcription initiation site (Tang *et al.* 1999). The PZ120 gene is now replaced in the NCBI database as the zinc finger and BTB domain containing 11 (ZBTB11). So far, no additional reports describing this transcription factor were published since. ETF is present in all MT isoforms, and has not been described in the context of MT inhibition. Overexpression of this factor in primate kidney CV1 cells showed that it represses expression originating from both the EGFR and beta-actin gene promoters (Kageyama and Pastan 1989).

The influence of environmental factors on MT expression was illustrated by the inhibition of zinc induced MT induction by chromium, a major environmental carcinogen (Majumder *et al.* 2003). This inhibition was working through interfering with MTF1.

Treatment of mammalian cells with cycloheximide, a protein synthesis inhibitor, resulted in increased MT1 transcription (McCormick *et al.* 1991). Recently, the presence of a labile inhibitor of MT1 expression was suggested (Bi *et al.* 2006). This repressor negatively controls agonist-induced turnover of the MTF1 protein.

Promoter methylation

Methylation of cytosine residues in promoter sequences is generally associated with a low transcriptional level of the respective gene. In mammals, methylation is mainly found within so-called CpG islands, regions of high CG content, thought to be involved in transcriptional regulation. The general rule is that CpG islands are not methylated, except for genes on the inactive X chromosome and at imprinted loci. Generally, to study the methylation status of a gene in a cell line or in tissue, its expression is correlated to the amount of methylated CG dinucleotides in genomic DNA isolated from the same source. If a correlation is found, e.g. low gene expression and a hypermethylation of the CpG island, the cells or tissue specimens are subjected to a demethylation agent, such as 5-azacytidine. If the demethylation results in an increased expression of the gene, a role for methylation in expression regulation of the gene is established.

Rodent and human MT genes contain a CpG island in their promoter. Tissue specific methylation of MT1B has been demonstrated (Heguy *et al.* 1986). The MT1B gene is only expressed in human hepatoma and renal carcinoma cell lines, and not in HeLa cells, where the 5' flanking region of MT1B is highly methylated. MT3 hypermethylation has been shown in gastric cancer (Deng *et al.* 2003) and in oesophagal squamous carcinoma (Smith *et al.* 2005).

CONCLUSION

The persistent differential expression of MTs in stress reaction and in pathological conditions suggests a strict regulation of these proteins. It is plausible that a change in e.g. cell cycle regulation of MT could influence cancer development. In addition, cell-type specific expression of MTs may indicate a divergence in functions within cell types or organs. For instance, disability of the tightly controlled proliferation of stem cells in the intestinal crypts may render individuals more susceptible for developing colon cancers. Similarly, a disturbed immune balance in the intestine due to changes in MT expression could be a prerequisite for IBD. Our *in silico* search for transcriptional regulatory regions in the MT promoters revealed new potential targets that could help to unravel some of the MT regulatory mechanisms. Future studies on the expression and regulation of MT genes are likely to provide insights to their role in both health and disease. Ultimately, novel strategies for manipulating intracellular MT levels could lead to new therapies.

REFERENCES

Adlard, P. A., A. K. West and J. C. Vickers (1998). "Increased density of metallothionein I/II-immunopositive cortical glial cells in the early stages of Alzheimer's disease." Neurobiol Dis **5**(5): 349-56.

Andrews, G. K., D. K. Lee, R. Ravindra*, et al.* (2001). "The transcription factors MTF-1 and USF1 cooperate to regulate mouse metallothionein-I expression in response to the essential metal zinc in visceral endoderm cells during early development." Embo J **20**(5): 1114-22.

Bankier, A. (1995). "Menkes disease." J Med Genet. **32**(3): 213-5.

Bi, Y., G. X. Lin, L. Millecchia*, et al.* (2006). "Superinduction of metallothionein I by inhibition of protein synthesis: Role of a labile repressor in MTF-1 mediated gene transcription." J Biochem Mol Toxicol **20**(2): 57-68.

Binz, P. A. a. K., J.H.R. (1999). Metallothionein: molecular evolution and classification. Metallothionein IV. , Klaassen, C.D., Ed., Basel, Birkhäuser Verlag**:** 7-13.

Blaauwgeers, H. G., M. Anwar Chand, F. M. van den Berg*, et al.* (1996). "Expression of different metallothionein messenger ribonucleic acids in motor cortex, spinal cord and liver from patients with amyotrophic lateral sclerosis." J Neurol Sci. **142**(1-2): 39-44.

Borghesi, L. A. and M. A. Lynes (1996). "Stress proteins as agents of immunological change: some lessons from metallothionein." Cell Stress Chaperones. **1**(2): 99-108.

Brugnera, E., O. Georgiev, F. Radtke*, et al.* (1994). "Cloning, chromosomal mapping and characterization of the human metal-regulatory transcription factor MTF-1." Nucleic Acids Res **22**(15): 3167-73.

Bruwer, M., K. W. Schmid, K. A. Metz*, et al.* (2001). "Increased expression of metallothionein in inflammatory bowel disease." Inflamm Res. **50**(6): 289-93.

Chelly, J., Z. Tumer, T. Tonnesen*, et al.* (1993). "Isolation of a candidate gene for Menkes disease that encodes a potential heavy metal binding protein." Nat Genet. **3**(1): 14-9.

Cherian, M. G. and M. D. Apostolova (2000). "Nuclear localization of metallothionein during cell proliferation and differentiation." Cell Mol Biol (Noisy-le-grand). **46**(2): 347-56.

Chung, R. S., P. A. Adlard, J. Dittmann*, et al.* (2004). "Neuron-glia communication: metallothionein expression is specifically up-regulated by astrocytes in response to neuronal injury." J Neurochem **88**(2): 454-61.

Clarkson, J. P., M. E. Elmes, B. Jasani*, et al.* (1985). "Histological demonstration of immunoreactive zinc metallothionein in liver and ileum of rat and man." Histochem J. **17**(3): 343-52.

Colangelo, V., J. Schurr, M. J. Ball*, et al.* (2002). "Gene expression profiling of 12633 genes in Alzheimer hippocampal CA1: transcription and neurotrophic factor down-regulation and up-regulation of apoptotic and pro-inflammatory signaling." J Neurosci Res **70**(3): 462-73.

Coto, J. A., E. M. Hadden, M. Sauro*, et al.* (1992). "Interleukin 1 regulates secretion of zinc-thymulin by human thymic epithelial cells and its action on T-lymphocyte proliferation and nuclear protein kinase C." Proc Natl Acad Sci U S A **89**(16): 7752-6.

Dalton, T., R. D. Palmiter and G. K. Andrews (1994). "Transcriptional induction of the mouse metallothionein-I gene in hydrogen peroxide-treated Hepa cells involves a composite major late transcription factor/antioxidant response element and metal response promoter elements." Nucleic Acids Res **22**(23): 5016-23.

Dameron, C. T. and M. D. Harrison (1998). "Mechanisms for protection against copper toxicity." Am J Clin Nutr. **67**(5 Suppl): 1091S-1097S.

Datta, P. K. and S. T. Jacob (1993). "Identification of a sequence within the mouse metallothionein-I gene promoter mediating its basal transcription and of a protein interacting with this element." Cell Mol Biol Res **39**(5): 439-49.

De Bleser, P., D. Vlieghe and F. Van Roy (2006). "A Distance Difference Matrix Approach to the Identification of Transcription Factors that Regulate Differential Gene Expression in Higher Eukaryotes." Genome Research.

De, S. K., M. T. McMaster and G. K. Andrews (1990). "Endotoxin induction of murine metallothionein gene expression." J Biol Chem **265**(25): 15267-74.

Deng, D., W. El-Rifai, J. Ji*, et al.* (2003). "Hypermethylation of metallothionein-3 CpG island in gastric carcinoma." Carcinogenesis **24**(1): 25-9.

Dooley, T. P., E. V. Curto, S. P. Reddy*, et al.* (2004). "Regulation of gene expression in inflammatory bowel disease and correlation with IBD drugs: screening by DNA microarrays." Inflamm Bowel Dis **10**(1): 1-14.

Duguid, J. R., C. W. Bohmont, N. G. Liu*, et al.* (1989). "Changes in brain gene expression shared by scrapie and Alzheimer disease." Proc Natl Acad Sci U S A **86**(18): 7260-4.

Dutta, R., D. A. Sens, S. Somji*, et al.* (2002). "Metallothionein isoform 3 expression inhibits cell growth and increases drug resistance of PC-3 prostate cancer cells
Metallothionein isoform 3 expression in the human prostate and cancer-derived cell lines." Prostate. **52**(2): 89-97.

Elmes, M. E., J. P. Clarkson and B. Jasani (1986). "The immunocytochemical demonstration of metallothionein in human liver and small intestine." Acta Pharmacol Toxicol (Copenh). **59**(Suppl 7): 510-3.

Erickson, J. C., A. K. Sewell, L. T. Jensen*, et al.* (1994). "Enhanced neurotrophic activity in Alzheimer's disease cortex is not associated with down-regulation of metallothionein-III (GIF)." Brain Res **649**(1-2): 297-304.

Falck, F. Y., Jr., L. J. Fine, R. G. Smith*, et al.* (1983). "Metallothionein and occupational exposure to cadmium." Br J Ind Med **40**(3): 305-13.

Fuller, C. E., M. E. Elmes and B. Jasani (1990). "Age-related changes in metallothionein, copper, copper-associated protein, and lipofuscin in human liver: a histochemical and immunohistochemical study." J Pathol **161**(2): 167-72.

Garnica, A. D., W. Y. Chan and O. M. Rennert (1978). "Role of metallothioneins in copper transport in patients with Menkes syndrome." Ann Clin Lab Sci. **8**(4): 302-9.

Gedamu, L., U. Varshney, N. Jahroudi*, et al.* (1987). "Structure and expression of the human metallothionein genes." Experientia Suppl **52**: 361-72.

Ghoshal, K. and S. T. Jacob (2001). "Regulation of metallothionein gene expression." Prog Nucleic Acid Res Mol Biol. **66**: 357-84.

Gong, Y. H. and J. L. Elliott (2000). "Metallothionein expression is altered in a transgenic murine model of familial amyotrophic lateral sclerosis." Exp Neurol. **162**(1): 27-36.

Haerslev, T., G. K. Jacobsen, N. Horn*, et al.* (1995). "Metallothionein expression in placental tissue in Menkes' disease. An immunohistochemical study." Apmis. **103**(7-8): 568-73.

Hamer, D. H. (1987). "Metallothionein gene regulation in Menkes' syndrome." Arch Dermatol. **123**(10): 1384a-1385a.

Heguy, A., A. West, R. I. Richards*, et al.* (1986). "Structure and tissue-specific expression of the human metallothionein IB gene." Mol Cell Biol. **6**(6): 2149-57.

Helin, K., C. L. Wu, A. R. Fattaey, *et al.* (1993). "Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperative trans-activation." Genes Dev **7**(10): 1850-61.

Hesketh, J. (2004). "3'-Untranslated regions are important in mRNA localization and translation: lessons from selenium and metallothionein." Biochem Soc Trans **32**(Pt 6): 990-3.

Heuchel, R., F. Radtke, O. Georgiev, *et al.* (1994). "The transcription factor MTF-1 is essential for basal and heavy metal-induced metallothionein gene expression." Embo J **13**(12): 2870-5.

Hoey, J. G., S. H. Garrett, M. A. Sens, *et al.* (1997). "Expression of MT-3 mRNA in human kidney, proximal tubule cell cultures, and renal cell carcinoma." Toxicol Lett. **92**(2): 149-60.

Hoogenraad, T. U. (2006). "Paradigm shift in treatment of Wilson's disease: Zinc therapy now treatment of choice." Brain Dev **5**: 5.

Inoue, K., H. Takano, A. Shimada, *et al.* (2006). "Role of metallothionein in coagulatory disturbance and systemic inflammation induced by lipopolysaccharide in mice." Faseb J **20**(3): 533-5.

Ioachim, E., M. Michael, C. Katsanos, *et al.* (2003). "The immunohistochemical expression of metallothionein in inflammatory bowel disease. Correlation with HLA-DR antigen expression, lymphocyte subpopulations and proliferation-associated indices." Histol Histopathol. **18**(1): 75-82.

Irie, Y., F. Mori, W. M. Keung, *et al.* (2004). "Expression of neuronal growth inhibitory factor (metallothionein-III) in the salivary gland." Physiol Res. **53**(6): 719-23.

Jacob, C., W. Maret and B. L. Vallee (1998). "Control of zinc transfer between thionein, metallothionein, and zinc proteins." Proc Natl Acad Sci U S A **95**(7): 3489-94.

Jacob, S. T., K. Ghoshal and J. F. Sheridan (1999). "Induction of metallothionein by stress and its molecular mechanisms." Gene Expr. **7**(4-6): 301-10.

Jahroudi, N., R. Foster, J. Price-Haughey, *et al.* (1990). "Cell-type specific and differential regulation of the human metallothionein genes. Correlation with DNA methylation and chromatin structure." J Biol Chem **265**(11): 6506-11.

Janssen, A. M., W. van Duijn, M. M. Oostendorp-Van De Ruit, *et al.* (2000). "Metallothionein in human gastrointestinal cancer." J Pathol **192**(3): 293-300.

Jasani, B., F. Campbell, H. Navabi, *et al.* (1998). "Clonal overexpression of metallothionein is induced by somatic mutation in morphologically normal colonic mucosa." J Pathol. **184**(2): 144-7.

Kageyama, R. and I. Pastan (1989). "Molecular cloning and characterization of a human DNA binding factor that represses transcription." Cell **59**(5): 815-25.

Kagi, J. H. and B. L. Valee (1960). "Metallothionein: a cadmium- and zinc-containing protein from equine renal cortex." J Biol Chem. **235**: 3460-5.

Kanekiyo, M., N. Itoh, A. Kawasaki*, et al.* (2002). "Metallothionein modulates lipopolysaccharide-stimulated tumour necrosis factor expression in mouse peritoneal macrophages." Biochem J **361**(Pt 2): 363-9.

Karin, M., R. L. Eddy, W. M. Henry*, et al.* (1984). "Human metallothionein genes are clustered on chromosome 16." Proc Natl Acad Sci U S A. **81**(17): 5494-8.

Karin, M. and H. R. Herschman (1979). "Dexamethasone stimulation of metallothionein synthesis in HeLa cell cultures." Science. **204**(4389): 176-7.

Kasutani, K., N. Itoh, M. Kanekiyo*, et al.* (1998). "Requirement for cooperative interaction of interleukin-6 responsive element type 2 and glucocorticoid responsive element in the synergistic activation of mouse metallothionein-I gene by interleukin-6 and glucocorticoid." Toxicol Appl Pharmacol **151**(1): 143-51.

Kelly, E. J., E. P. Sandgren, R. L. Brinster*, et al.* (1997). "A pair of adjacent glucocorticoid response elements regulate expression of two mouse metallothionein genes." Proc Natl Acad Sci U S A. **94**(19): 10045-50.

Kim, C. H., J. H. Kim, J. Lee*, et al.* (2003). "Zinc-induced NF-kappaB inhibition can be modulated by changes in the intracellular metallothionein level." Toxicol Appl Pharmacol **190**(2): 189-96.

Kodama, H. and Y. Murata (1999). "Molecular genetics and pathophysiology of Menkes disease." Pediatr Int. **41**(4): 430-5.

Kruidenier, L., I. Kuiper, W. Van Duijn*, et al.* (2003). "Imbalanced secondary mucosal antioxidant response in inflammatory bowel disease." J Pathol. **201**(1): 17-27.

Laukens, D., H. Peeters, S. Bogaert*, et al.* (2006). "Reduced metallothionein expression in colonic Crohn's disease: evidence for MTF1 as a new disease-modifying gene." *Submitted to Gastroenterology*.

Lawrance, I. C., C. Fiocchi and S. Chakravarti (2001). "Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes." Hum Mol Genet. **10**(5): 445-56.

Leibbrandt, M. E., R. Khokha and J. Koropatnick (1994). "Antisense down-regulation of metallothionein in a human monocytic cell line alters adherence, invasion, and the respiratory burst." Cell Growth Differ **5**(1): 17-25.

Majumder, S., K. Ghoshal, R. M. Gronostajski*, et al.* (2001). "Downregulation of constitutive and heavy metal-induced metallothionein-I expression by nuclear factor I." Gene Expr **9**(4-5): 203-15.

Majumder, S., K. Ghoshal, D. Summers*, et al.* (2003). "Chromium(VI) down-regulates heavy metal-induced metallothionein gene transcription by modifying transactivation potential

of the key transcription factor, metal-responsive transcription factor 1." J Biol Chem **278**(28): 26216-26.

Manuel, Y., Y. Thomas and O. Pellegrini (1992). "Metallothionein and tissue damage." IARC Sci Publ.(118): 231-7.

Martin, J., C. Han, L. A. Gordon*, et al.* (2004). "The sequence and analysis of duplication-rich human chromosome 16." Nature. **432**(7020): 988-94.

Mathew, C. G. and C. M. Lewis (2004). "Genetics of inflammatory bowel disease: progress and prospects." Hum Mol Genet **13 Spec No 1**: R161-8.

McCormick, C. C., L. M. Salati and A. G. Goodridge (1991). "Abundance of hepatic metallothionein mRNA is increased by protein-synthesis inhibitors. Evidence for transcriptional activation and post-transcriptional regulation." Biochem J **273(Pt 1)**: 185-8.

Meplan, C., K. Mann and P. Hainaut (1999). "Cadmium induces conformational modifications of wild-type p53 and suppresses p53 response to DNA damage in cultured cells." J Biol Chem. **274**(44): 31663-70.

Mercer, J. F., J. Livingston, B. Hall*, et al.* (1993). "Isolation of a partial candidate gene for Menkes disease by positional cloning." Nat Genet. **3**(1): 20-5.

Mitropoulos, D., A. Kyroudi-Voulgari, S. Theocharis*, et al.* (2005). "Prognostic significance of metallothionein expression in renal cell carcinoma." World J Surg Oncol. **3**(1): 5.

Miyashita, H. and Y. Sato (2005). "Metallothionein 1 is a downstream target of vascular endothelial zinc finger 1 (VEZF1) in endothelial cells and participates in the regulation of angiogenesis." Endothelium. **12**(4): 163-70.

Moffatt, P. and C. Seguin (1998). "Expression of the gene encoding metallothionein-3 in organs of the reproductive system." DNA Cell Biol. **17**(6): 501-10.

Morahan, J. M., B. Yu, R. J. Trent*, et al.* (2005). "Screening the metallothionein III gene in sporadic amyotrophic lateral sclerosis." Amyotroph Lateral Scler Other Motor Neuron Disord. **6**(2): 115-7.

Mulder, T. P., A. van der Sluys Veer, H. W. Verspaget*, et al.* (1994). "Effect of oral zinc supplementation on metallothionein and superoxide dismutase concentrations in patients with inflammatory bowel disease." J Gastroenterol Hepatol **9**(5): 472-7.

Mulder, T. P., H. W. Verspaget, A. R. Janssens*, et al.* (1991). "Decrease in two intestinal copper/zinc containing proteins with antioxidant function in inflammatory bowel disease." Gut. **32**(10): 1146-50.

Nagano, S., M. Satoh, H. Sumi*, et al.* (2001). "Reduction of metallothioneins promotes the disease expression of familial amyotrophic lateral sclerosis mice in a dose-dependent manner." Eur J Neurosci. **13**(7): 1363-70.

Nagel, W. W. and B. L. Vallee (1995). "Cell cycle regulation of metallothionein in human colonic cancer cells." Proc Natl Acad Sci U S A. **92**(2): 579-83.

Nury, D., H. Chabanon, M. Levadoux-Martin*, et al.* (2005). "An eleven nucleotide section of the 3'-untranslated region is required for perinuclear localization of rat metallothionein-1 mRNA." Biochem J **387**(Pt 2): 419-28.

Ostrakhovitch, E. A., P. E. Olsson, S. Jiang*, et al.* (2006). "Interaction of metallothionein with tumor suppressor p53 protein." FEBS Lett **580**(5): 1235-8.

Oz, H. S., T. Chen, W. J. de Villiers*, et al.* (2005). "Metallothionein overexpression does not protect against inflammatory bowel disease in a murine colitis model." Med Sci Monit **11**(3): BR69-73.

Penkowa, M. (2006). "Metallothioneins are multipurpose neuroprotectants during brain pathology." FEBS Journal **273**(9): 1857-1870.

Quaife, C. J., S. D. Findley, J. C. Erickson*, et al.* (1994). "Induction of a new metallothionein isoform (MT-IV) occurs during differentiation of stratified squamous epithelia." Biochemistry. **33**(23): 7250-9.

Roy, A. L., M. Meisterernst, P. Pognonec*, et al.* (1991). "Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF." Nature **354**(6350): 245-8.

Sakurai, A., S. Hara, N. Okano*, et al.* (1999). "Regulatory role of metallothionein in NF-kappaB activation." FEBS Lett **455**(1-2): 55-8.

Saydam, N., T. K. Adams, F. Steiner*, et al.* (2002). "Regulation of metallothionein transcription by the metal-responsive transcription factor MTF-1: identification of signal transduction cascades that control metal-inducible transcription." J Biol Chem **277**(23): 20438-45.

Schmidt, C. J. and D. H. Hamer (1986). "Cell specificity and an effect of ras on human metallothionein gene expression." Proc Natl Acad Sci U S A. **83**(10): 3346-50.

Schmidt, C. J., D. H. Hamer and O. W. McBride (1984). "Chromosomal location of human metallothionein genes: implications for Menkes' disease." Science. **224**(4653): 1104-6.

Schmidt, C. J., M. F. Jubier and D. H. Hamer (1985). "Structure and expression of two human metallothionein-I isoform genes and a related pseudogene." J Biol Chem. **260**(12): 7731-7.

Schroeder, J. J. and R. J. Cousins (1990). "Interleukin 6 regulates metallothionein gene expression and zinc metabolism in hepatocyte monolayer cultures." Proc Natl Acad Sci U S A **87**(8): 3137-41.

Schwede, T., J. Kopp, N. Guex*, et al.* (2003). "SWISS-MODEL: An automated protein homology-modeling server." Nucleic Acids Res **31**(13): 3381-5.

Searle, P. F., B. L. Davison, G. W. Stuart*, et al.* (1984). "Regulation, linkage, and sequence of mouse metallothionein I and II genes." Mol Cell Biol. **4**(7): 1221-30.

Shworak, N. W., T. O'Connor, N. C. Wong*, et al.* (1993). "Distinct TATA motifs regulate differential expression of human metallothionein I genes MT-IF and MT-IG." J Biol Chem **268**(32): 24460-6.

Sillevis Smitt, P. A., H. G. Blaauwgeers, D. Troost*, et al.* (1992a). "Metallothionein immunoreactivity is increased in the spinal cord of patients with amyotrophic lateral sclerosis." Neurosci Lett. **144**(1-2): 107-10.

Sillevis Smitt, P. A., T. P. Mulder, H. W. Verspaget*, et al.* (1994). "Metallothionein in amyotrophic lateral sclerosis." Biol Signals. **3**(4): 193-7.

Sillevis Smitt, P. A., H. van Beek, A. J. Baars*, et al.* (1992b). "Increased metallothionein in the liver and kidney of patients with amyotrophic lateral sclerosis." Arch Neurol. **49**(7): 721-4.

Smith, E., P. A. Drew, Z. Q. Tian*, et al.* (2005). "Metallothionien 3 expression is frequently down-regulated in oesophageal squamous cell carcinoma by DNA methylation." Mol Cancer **4**: 42.

Stefanidou, M., C. Maravelias, A. Dona*, et al.* (2006). "Zinc: a multipurpose trace element." Arch Toxicol **80**(1): 1-9.

Studer, R., C. P. Vogt, M. Cavigelli*, et al.* (1997). "Metallothionein accretion in human hepatic cells is linked to cellular proliferation." Biochem J. **328**(Pt 1): 63-7.

Sturniolo, G. C., C. Mestriner, P. E. Lecis*, et al.* (1998). "Altered plasma and mucosal concentrations of trace elements and antioxidants in active ulcerative colitis." Scand J Gastroenterol. **33**(6): 644-9.

Szczurek, E. I., C. S. Bjornsson and C. G. Taylor (2001). "Dietary zinc deficiency and repletion modulate metallothionein immunolocalization and concentration in small intestine and liver of rats." J Nutr **131**(8): 2132-8.

Takeda, A., H. Hisada, S. Okada*, et al.* (1997). "Tumor cell growth is inhibited by suppressing metallothionein-I synthesis." Cancer Lett. **116**(2): 145-9.

Tang, C. M., J. Westling and E. Seto (1999). "trans repression of the human metallothionein IIA gene promoter by PZ120, a novel 120-kilodalton zinc finger protein." Mol Cell Biol **19**(1): 680-9.

Tate, D. J., M. V. Miceli and D. A. Newsome (2002). "Expression of metallothionein isoforms in human chorioretinal complex." Curr Eye Res. **24**(1): 12-25.

Theocharis, S. E., A. P. Margeli, J. T. Klijanienko*, et al.* (2004). "Metallothionein expression in human neoplasia." Histopathology **45**(2): 103-18.

Tomita, T. and O. Matsubara (2000). "Immunocytochemical localization of metallothionein in human pancreatic islets." Pancreas **20**(1): 21-4.

Tran, C. D., R. N. Butler, G. S. Howarth*, et al.* (1999). "Regional distribution and localization of zinc and metallothionein in the intestine of rats fed diets differing in zinc content." Scand J Gastroenterol **34**(7): 689-95.

Tsuji, S., H. Kobayashi, Y. Uchida*, et al.* (1992a). "Molecular cloning of human growth inhibitory factor cDNA and its down-regulation in Alzheimer's disease." Embo J. **11**(13): 4843-50.

Tsuji, S., H. Kobayashi, Y. Uchida*, et al.* (1992b). "Molecular cloning of human growth inhibitory factor cDNA and its down-regulation in Alzheimer's disease." Embo J **11**(13): 4843-50.

van den Oord, J. J. and M. De Ley (1994). "Distribution of metallothionein in normal and pathological human skin." Arch Dermatol Res. **286**(1): 62-8.

Varshney, U., N. Jahroudi, R. Foster*, et al.* (1986). "Structure, organization, and regulation of human metallothionein IF gene: differential and cell-type-specific expression in response to heavy metals and glucocorticoids." Mol Cell Biol **6**(1): 26-37.

Vulpe, C., B. Levinson, S. Whitney*, et al.* (1993). "Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase." Nat Genet. **3**(1): 7-13.

Yin, X., D. A. Knecht and M. A. Lynes (2005). "Metallothionein mediates leukocyte chemotaxis." BMC Immunol **6**: 21.

Yu, W. H., W. J. Lukiw, C. Bergeron*, et al.* (2001). "Metallothionein III is reduced in Alzheimer's disease." Brain Res **894**(1): 37-45.

Zambenedetti, P., R. Giordano and P. Zatta (1998). "Metallothioneins are highly expressed in astrocytes and microcapillaries in Alzheimer's disease." J Chem Neuroanat **15**(1): 21-6.

-Section II-

Chapter 5

SPONDYLOARTHROPATHY AS A MODEL FOR EARLY CROHN'S DISEASE

Debby Laukens

Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium

INTRODUCTION

The spondyloarthropathies (SpA) are a heterogeneous group of chronic inflammatory arthritides that share certain clinical features, and are linked by their association to the human leukocyte antigen class-I gene, *HLA-B27* (Wright 1978). The most characteristic clinical feature is inflammatory back pain, caused by inflammation of the sacroiliac joints (sacroiliitis, Figure 1 left), or the joints of the spinal vertebrae (spondylitis, Figure 1 middle). A specific characteristic of SpA is enthesitis, involving inflammation at the sites where tendons, ligaments or joint capsules are attached to the bone (Figure 1, right). Peripheral joint inflammation is frequently present in SpA, and is mostly pauciarticular (less than four joints), and affects predominantly the lower limbs. These are progressive diseases in which chronic inflammation leads to deterioration of the bone or cartilage. New bone formation in the spine or peripheral joints severely impairs movement in these patients. Usually, low back pain is the first symptom, which is worse at night, in the morning and after periods of inactivity. The course of SpA is highly variable and characterized by spontaneous remissions and flare-ups, particularly in the early stages. Often, extra-articular manifestations are associated with SpA, such as uveitis, psoriasis and inflammatory bowel disease (IBD), either Crohn's disease (CD) or ulcerative colitis (UC).
Since 1991, the European Spondyloarthropathy Study Group (ESSG) criteria are widely used for diagnosis of SpA (Dougados *et al.* 1991). The SpA group is generally divided into 5 main disease categories (Table 1): 1) ankylosing spondylitis (AS), the most common form of SpA; 2)

Table 1 Features of spondyloarthropathies

	Anklosing spondylitis	**Reactive arthritis**	**Psoriatic arthritis**	**SpA associated with IBD**
features				
prevalence	0.1-0.2%	0.1%	0.2-0.4%	rare
age at onset	late teens to early adulthood	late teens to early adulthood	35 to 45 years	any age
male to female ratio	3:1	5:1	1:1	1:1
HLA-B27	90-95%	80%	40%	30%
articular manifestations				
sacroiliitis				
frequency	100%	40-60%	40%	20%
distribution	symmetric	asymmetric	asymmetric	asymmetric
peripheral arthritis				
frequency	occasional	common	common	common
distribution	asymmetric, lower limbs	asymmetric, lower limbs	asymmetric, any joint	asymmetric, lower limbs
enthesitis	common	very common	very common	occasional
dactylitis	uncommon	common	common	uncommon
syndesmophutes	delicate, marginal	bulky, nonmarginal	bulky, nonmarginal	delicate, marginal
extra-articular manifestations				
skin lesions	none	circinate balanitis, keratoderma, blennorrhagicum	psoriasis	erythema nodosum, pyroderma gangrenosum
nail changes	none	onycholysis	pitting, onycholysis	clubbing
ocular conditions	acute anterior uveitis	acute anterior uveitis, conjunctivitis	chronic uveitis	chronic uveitis
cardiac conditions	aortic regurgitation, conduction defects	aortic regurgitation, conduction defects	aortic regurgitation, conduction defects	aortic regurgitation
oral conditions	ulcers	ulcers	ulcers	ulcers
pulmonary features	upper lobe fibrosis	none	none	none
gastrointestinal conditions	none	diarrhea	none	inflammatory bowel disease
genitorinary conditions	prostatitis	urethritis, cervicitis	none	none
renal conditions	amyloidosis, IgA nephropathy	amyloidosis	amyloidosis	nephrolithiasis

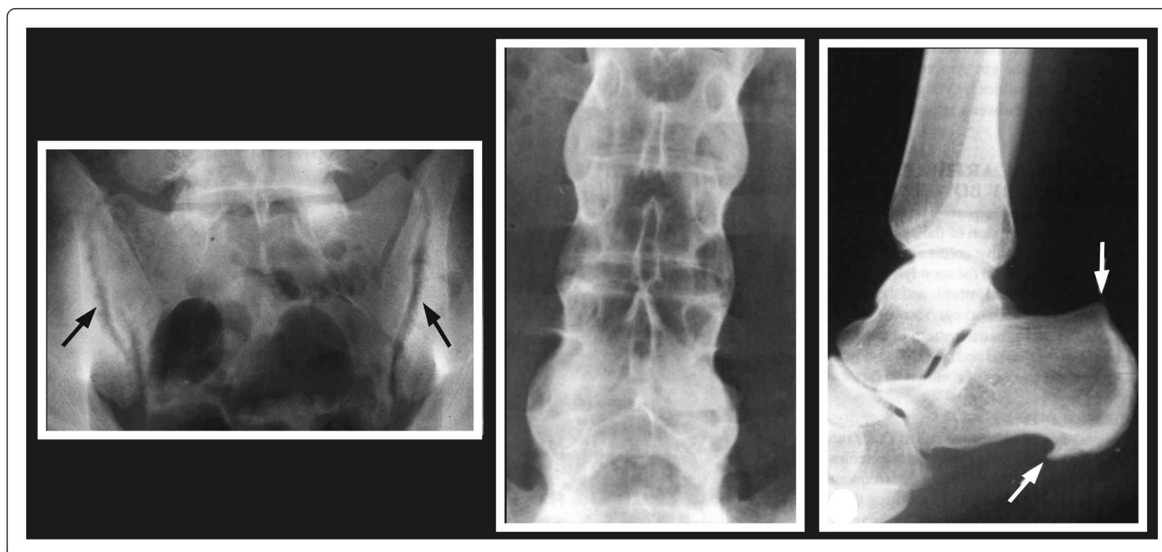


Figure 1 Radiographs of (left) the pelvis in a patient with sacroiliitis showing sclerosis and erosion of the sacroiliac joints (arrows); (middle) the lumbar spine in a patient with ankylosing spondylitis with complete ossification of the annulus fibrosus; (right) the heel in a patient with periosteal reaction at the plantar fascia insertion and at the Achilles tendon insertion on the calcaneus.

reactive arthritis (ReA), in patients with a recent history of a urogenital or an intestinal infection; 3) psoriatic arthritis (PsA); 4) SpA associated with IBD (SpA-IBD); and 5) undifferentiated spondyloarthropathy (USpA), grouping patients that fulfil the criteria for SpA but are not classifiable in one of the former groups.

There is an important overlap between the different disease manifestations of SpA in families. This argues for a genetic predisposition for SpA (Breban *et al.* 2003). As mentioned above, the SpAs are linked by a common genetic risk factor, *HLA-B27*, which is much more prevalent in SpA patients than in other rheumatic diseases or healthy controls (Table 1). HLA class I molecules are highly polymorphic membrane glycoproteins, and are specialized antigen-presenting molecules that form stable complexes with antigenic peptides, displaying them for recognition by $CD8^{+}$ T cells. It is anchored to the cell membrane by a short transmembrane segment, and consists of three α domains, non-covalently bounded to β2-microglobulin. There are currently 24 genotypic subtypes of *HLA-B27* identified, and *HLA-B*2705* bears the strongest association to SpA. *HLA-B*2702, *2703, *2704*, and **2707* are associated with AS. Worth mentioning is that the prevalence of SpA is correlated to the occurrence of *HLA-B27*. For example, SpA is frequent among Eskimo's, where the prevalence of *HLA-27* is 25 to 40% (Boyer *et al.* 1994). Conversely, SpA is rare in the Japanese population, where *HLA-B27* prevalence is less than 1% (Hukuda *et al.* 2001). The mechanism by which *HLA- B27* confers susceptibility to SpA is not understood. Several hypothesis have been proposed (Kim *et al.*

2005), including the arthritogenic peptide theory, molecular mimicry and aberrant processing or folding of the heavy chain of HLA-B27.

Although SpA is regarded as one entity of inflammatory arthritides, they display distinct clinical features. It is likely that interplay among genetic and environmental factors is responsible for the various clinical manifestations. Spondyloarthropathy is considered to result from infection or exposure to an unknown antigen in genetically susceptible patients. In the case of ReA, a known infection precedes the arthritis by several weeks. This arthritis is thought to be a post-infective phenomenon, rather than resulting from a direct infection, since up to now, no viable bacteria could be cultured from synovial fluid of inflamed joints. This phenomenon is also referred to as a sterile inflammation of the joint, as compared to infectious arthritis caused by direct invasion of the joint space by various micro-organisms. Nevertheless, bacterial antigens were commonly found in the synovium of SpA patients (see next paragraph). Thus, it can never be ruled out that an actual infection has occurred in the patient before the onset of SpA. It has been suggested that the gut is an important portal for antigenic uptake in SpA, and thus plays a pathogenic role in susceptible hosts. Here, we review this remarkable link between the bowel and the joint in SpA.

INTESTINAL INFLAMMATION, SPONDYLOARTHROPATHY AND IBD

An interesting link has been found between SpA, intestinal inflammation and IBD (Mielants *et al.* 2005). In 10 to 20% of IBD patients, inflammatory peripheral arthropathy has been observed, and 7 to 25% have axial involvement (de Vlam *et al.* 2000; De Vos 2004). Moreover, radiographic sacroiliitis is present in 20 to 25% of IBD patients. Clinically, these forms of arthritis are almost identical to SpA, although they are not associated with *HLA-B27*, indicating a different genetic predisposition. On the other hand, the gut is an important site of inflammation in patients with SpA. In ileocolonoscopic studies of SpA patients, histological signs of gut inflammation were found in more than half of the patients, mostly not presenting any clinical intestinal manifestations, while they were not seen in any other inflammatory joint disease (De Keyser *et al.* 1998). Remission of the joint inflammation was always linked with a disappearance of the gut inflammation. Two types of inflammation were distinguished: acute inflammation resembling infectious enterocolitis, and chronic inflammation more suggestive of early CD (Cuvelier *et al.* 1987). In the acute type of inflammation, the mucosal architecture is well preserved. The ileal villi and crypts are infiltrated by polymorphonuclear cells, while in the lamina propria, there is an increased number of inflammatory cells. The chronic type of inflammation is characterized by a clearly disturbed mucosal architecture, and is mostly

indistinguishable from CD. The villi are blunted and fused. The crypts are distorted and the lamina propria is edematous and infiltrated by mononuclear cells. Basal lymphoid follicles occur. In some cases, aphtoid ulcers and granulomas are present. The clinical significance of CD-like alterations in the bowel of SpA patients was shown in a prospective long-term study in which SpA patients were reviewed after several years (Mielants *et al.* 1995; De Vos *et al.* 1996). About 13% of patients who showed subclinical chronic gut inflammation on biopsies at the first investigation, developed full-blown CD. This supports the concept of subclinical CD in a subpopulation of SpA patients. Hence, SpA and CD could be regarded as a result of a common inflammatory pathway leading to phenotypes with comparable clinical and pathogenic features.
A number of studies have been performed to provide molecular and genetic support for the clinical observation of the association between gut and joint inflammation in SpA. These are discussed in the next paragraphs.

Genetic arguments for a joint-gut axis

Several studies indicated the specific genetic contribution of intestinal inflammation in SpA patients. Bjarnason and co-workers assessed the presence and inheritance pattern of subclinical intestinal inflammation in first-degree relatives of patients with AS (Bjarnason *et al.* 2003). They appeared to have an inherited abnormality that leads to subclinical intestinal inflammation, suggesting that this feature is transmitted and thus genetically determined. Detailed genetic analysis to identify risk factors for intestinal inflammation still needs to be done.
We recently performed a prospective clinical and radiological evaluation of 102 CD patients, and found an association between CD-associated *CARD15* polymorphisms (Chapter 1) (Hugot *et al.* 2001; Ogura *et al.* 2001) and the presence of sacroiliitis: 78% of patients with sacroiliitis were carriers of a mutation in *CARD15*, compared to 48% of patients without sacroiliitis (Peeters *et al.* 2004). Here, the carriage of *CARD15* mutations is predisposing for the onset of both chronic gut inflammation and sacroiliitis. The molecular mechanism of this association has yet to be investigated. However, replication of these results failed in a multicentre study, so caution must be taken in interpreting the involvement of *CARD15* in sacroiliitis.
The most convincing evidence for a joint-gut axis in SpA comes from the *HLA-B27* transgenic rats. These animals develop an illness similar to SpA, with manifestations including sacroiliitis, enthesitis, arthritis, skin and nail lesions, ocular inflammation, cardiac inflammation, and inflammation of the gastrointestinal and male genitourinary tracts (Taurog *et al.* 1994). Furthermore, the severity of the clinical disease correlates with the number of copies of *HLA-B27* expressed in the transgenic animal. This model provides direct evidence that over-

expression of the major risk factor in human SpA is causative for both arthritis and bowel inflammation, and supports the existence of an interplay between the gut and the joints. Moreover, if these *HLA-B27* transgenic rats are raised in a germ-free environment, they do not develop clinical disease. Once introduced to a regular environment and exposed to bacteria, the rats develop clinical manifestations of SpA, suggesting that commensal flora and/or pathogens are required in the disease onset. In humans, *HLA-B27* is strongly associated with AS and ReA (Table 1), and it is plausible that this protein can be pathogenically related to the rat model. Nevertheless, SpA with associated IBD is not highly correlated to *HLA-B27*: only 30% of this group of patients is positive for *HLA-B27*. In fact, of the *HLA-B27* negative SpA patients, ReA or USpA was diagnosed (Mielants *et al.* 1993). One explanation might be that there are two forms of SpA related with IBD, one in which *HLA-B27* is causative for both joint and gut inflammation, and one where another risk factor is involved. Genetically, it is possible that polymorphisms in another gene, located near *HLA-B27*, are associated with SpA, as it is believed that transmission of genetic information is confined to haplotype blocks, segments of DNA in which the chance of recombination is very low (Daly *et al.* 2001). On the other hand, the presence of HLA-B27 proteins might not be sufficient to disturb both joint and gut inflammatory pathways, a feature that is maybe not involved in rodents. Moreover, because the frequency of *HLA-B27* in the population-at-large far exceeds that of SpA, suggests that there are more genetic determinants for SpA (Reveille 2004). The search for new genes that are both involved in IBD and SpA is a challenging prospect for the future.

Immunologic arguments for a joint-gut axis

The immune infiltrate in the gut mucosa of patients with SpA is significantly different from that of healthy controls, even in the absence of inflammation, suggesting the presence of early inflammatory changes in the intestine of SpA patients. Lymphoid follicles in the intestine are specialized microenvironments where naive T cells respond to specific antigens encountered in the intestinal lumen. Under normal conditions, the number of follicles is limited. However, in the presence of inflammatory stimuli, their number increases. Similar to CD, the number of follicles is increased in the ileum and colon of SpA patients with microscopical normal gut biopsies (Demetter *et al.* 2002). The amount of leukocytes expressing CD11c was augmented in the ileum of SpA patients, while in the colon CD11a and VCAM-1 expressing leukocytes and $CD68^+$ macrophages were higher as compared to controls. The increase in $CD68^+$ macrophages and adhesion molecules was also seen in CD (Bernstein *et al.* 1998). Moreover, the expression of $\alpha E\beta 7$ integrin, a specific gut homing receptor for effector T cells, was up-regulated on gut

mucosal T cell lines isolated from SpA patients in the absence of inflammation (Elewaut *et al.* 1999; Van Damme *et al.* 2001c). In addition, an up-regulation of E-cadherin, the ligand for αEβ7, and its associated catenins, crucial for intercellular adhesion in epithelial cells, has been demonstrated in clinically overt IBD (Demetter *et al.* 2000). These observations are indicative of an increased infiltration of T cells and macrophages in the epithelial lining of the gut of SpA patients. Macrophages carrying the scavenger receptor CD163 were increased in colonic mucosa of CD and SpA patients, but not in UC (Demetter *et al.* 2005). They were also increased in the synovium of SpA patients, where they correlated with inflammatory parameters (Baeten *et al.* 2002). These $CD163^+$ macrophages produce interleukin 1 (IL1) and tumor necrosis factor alpha (TNF), but not IL10 following lipopolysaccharide (LPS) challenge (Baeten *et al.* 2002). Thus, an imbalance in cytokine production might lead to chronic inflammatory process in the gut mucosa of SpA and CD patients. Finally, there is a similar predominance of T helper type 1 (Th1) producing mucosal T cells in both SpA and CD, most surprisingly with a proportional decrease of interferon gamma and IL2 producing lymphocytes (Van Damme *et al.* 2001a; Van Damme *et al.* 2001b).

Overall, these observations point to the presence of early immune changes in the gut of patients with SpA, and an increased antigen handling and presentation, comparable to changes seen in CD.

Therapeutical intervention is effective for treatment of both joint and gut inflammation

Therapeutic intervention in SpA and CD can have a positive influence on both the articular and the intestinal inflammation. Sulfasalazine, which has been successfully used to treat colonic inflammation in UC and CD, was effective in the treatment of the peripheral arthritis accompanying SpA, especially if intestinal inflammation is present (Mielants *et al.* 1996).

Therapy based on a chimeric monoclonal antibody to TNF, infliximab, was highly effective for the treatment of CD. A pilot study conducted in four patients with CD as well as SpA showed a remission of gut inflammation and a significant improvement of articular and axial symptoms (Van den Bosch *et al.* 2000). Based on these initial findings, the use of infliximab was explored in a number of studies in patients with different forms of active SpA, and high success rates were reported. Based on these findings, two double-blind, placebo controlled trials were conducted (Braun *et al.* 2002; Van Den Bosch *et al.* 2002). A fast and significant improvement of disease was shown in these patients.

Although therapeutical improvement of both joint and gut inflammation is not a valid proof of a direct link between the two sites, it is however a nice example of how clinical observations led to

a significant improvement of treatment in these patients.

MECHANISM OF JOINT-GUT AXIS

The exact mechanism that links gut inflammation and joint inflammation is not completely understood. The main hypothesis states that exogenous factors, probably bacterial antigens, are permitted to enter the body through the gut. Bacterial antigens and LPS have been found in the synovial fluid of patients with ReA (Granfors *et al.* 1989; Granfors *et al.* 1990). In addition, evidence for the presence of bacterial DNA was shown by PCR in synovial fluid of patients with SpA (Pacheco-Tena *et al.* 2001). The delivery of bacterial antigens to the joint may trigger and perpetuate local inflammation, and is consistent with the parallelism of flare-ups of joint and intestinal inflammation. How these factors reach the joints is not known, given the fact that no live bacteria can be found. A role for T cells, HLA-B27 and macrophages has been postulated.

Increased priming of T cells in the SpA gut fits with the concept that intestinal T cells are involved in the induction of arthritis in ReA (Gaston 1993). Arthritis might arise from a T cell mediated immune response to bacterial antigens and degradation products circulating from the gut to the joint. A proof of concept for the T cell re-circulation hypothesis was put forward by the identification of similar clonal T cell expansion in the colonic mucosa and synovium of a patient with enterogenic SpA (May *et al.* 2000). In addition, *Yersinia* specific antigenic proliferation of T cell clones, isolated from the synovial fluid of a *Yersinia* triggered ReA patients has been shown (Hermann *et al.* 1989). This could mean that the same antigenic compounds are present in the joints and the gut of these patients. On the other hand, it was shown that gut-derived leukocytes from patients with IBD bind well to venules in the synovial membrane (Salmi *et al.* 2001), suggesting that intestinally activated T cells have the capacity to enter the synovium. So far, it is not clear whether the presence of similar T cell clones in joint and gut is due to the presence of identical antigens or because the same set of homing molecules occur at both sites. One can speculate that if similar antigenic peptides are present in the gut and the joint, a sustained HLA-B27 mediated T cell response can trigger inflammation.

Macrophages play a central role in the innate immune recognition of bacterial products. The finding that a similar subset of $CD163^+$ macrophages was found in both the synovium and gut mucosa of SpA patients could reflect a similar pathophysiological condition leading to the influx or maturation of this particular macrophage subset at the two sites. It is surprising that stimulation of these cells with LPS does not induce IL10 secretion, while stimulation with the natural CD163 ligand, haptoglobin-hemoglobin complexes, does actually lead to an anti-

inflammatory response of these cells (Ugocsai *et al.* 2006). It seems plausible that these cells may induce a similar dysregulation of the cytokine imbalance in the colon and synovium.

CONCLUSION

The ample clinical, immunological and genetic data support the existence of a joint-gut connection, explaining the important role of the gut in the pathogenesis of SpA. Maybe, these observations must not be separated from the involvement of other mucosal sites within the concept of SpA, e.g. uveitis and psoriasis. A similar mechanism might be involved in these clinical manifestations. Ultimately, knowledge of molecular events in these processes is important for classification, diagnosis and therapy of SpA patients.

REFERENCES

Baeten, D., P. Demetter, C. A. Cuvelier, *et al.* (2002). "Macrophages expressing the scavenger receptor CD163: a link between immune alterations of the gut and synovial inflammation in spondyloarthropathy." J Pathol. **196**(3): 343-50.

Bernstein, C. N., M. Sargent and W. M. Gallatin (1998). "Beta2 integrin/ICAM expression in Crohn's disease." Clin Immunol Immunopathol. **86**(2): 147-60.

Bjarnason, I., K. O. Helgason, A. J. Geirsson, *et al.* (2003). "Subclinical intestinal inflammation and sacroiliac changes in relatives of patients with ankylosing spondylitis." Gastroenterology **125**(6): 1598-605.

Boyer, G. S., D. W. Templin, J. C. Cornoni-Huntley, *et al.* (1994). "Prevalence of spondyloarthropathies in Alaskan Eskimos." J Rheumatol. **21**(12): 2292-7.

Braun, J., J. Brandt, J. Listing, *et al.* (2002). "Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial." Lancet. **359**(9313): 1187-93.

Breban, M., R. Said-Nahal, J. P. Hugot, *et al.* (2003). "Familial and genetic aspects of spondyloarthropathy." Rheum Dis Clin North Am. **29**(3): 575-94.

Cuvelier, C., C. Barbatis, H. Mielants, *et al.* (1987). "Histopathology of intestinal inflammation related to reactive arthritis." Gut **28**(4): 394-401.

Daly, M. J., J. D. Rioux, S. F. Schaffner, *et al.* (2001). "High-resolution haplotype structure in the human genome." Nat Genet. **29**(2): 229-32.

De Keyser, F., D. Elewaut, M. De Vos, *et al.* (1998). "Bowel inflammation and the spondyloarthropathies." Rheum Dis Clin North Am. **24**(4): 785-813.

de Vlam, K., H. Mielants, C. Cuvelier, *et al.* (2000). "Spondyloarthropathy is underestimated in inflammatory bowel disease: prevalence and HLA association." J Rheumatol. **27**(12):

2860-5.

De Vos, M., H. Mielants, C. Cuvelier*, et al.* (1996). "Long-term evolution of gut inflammation in patients with spondyloarthropathy." Gastroenterology **110**(6): 1696-703.

De Vos, M. (2004). "Review article: joint involvement in inflammatory bowel disease." Aliment Pharmacol Ther. **20**(Suppl 4): 36-42.

Demetter, P., M. De Vos, N. Van Damme*, et al.* (2000). "Focal up-regulation of E-cadherin-catenin complex in inflamed bowel mucosa but reduced expression in ulcer-associated cell lineage." Am J Clin Pathol. **114**(3): 364-70.

Demetter, P., J. A. Van Huysse, F. De Keyser*, et al.* (2002). "Increase in lymphoid follicles and leukocyte adhesion molecules emphasizes a role for the gut in spondyloarthropathy pathogenesis." J Pathol **198**(4): 517-22.

Demetter, P., M. De Vos, J. A. Van Huysse*, et al.* (2005). "Colon mucosa of patients both with spondyloarthritis and Crohn's disease is enriched with macrophages expressing the scavenger receptor CD163." Ann Rheum Dis. **64**(2): 321-4.

Dougados, M., S. van der Linden, R. Juhlin*, et al.* (1991). "The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy." Arthritis Rheum **34**(10): 1218-27.

Eckmann, L. and M. Karin (2005). "NOD2 and Crohn's disease: loss or gain of function?" Immunity. **22**(6): 661-7.

Elewaut, D., N. Van Damme, D. Baeten*, et al.* (1999). "Intestinal mucosa of patients with spondyloarthropathy is enriched with T cells carrying αEβ7 integrin, even in the absence of histologically defined inflammation." Gastroenterology **116**: G3069.

Gaston, H. (1993). "Synovial lymphocytes and the aetiology of synovitis." Ann Rheum Dis. **52**(Suppl 1): S17-21.

Granfors, K., S. Jalkanen, R. von Essen*, et al.* (1989). "Yersinia antigens in synovial-fluid cells from patients with reactive arthritis." N Engl J Med. **320**(4): 216-21.

Granfors, K., S. Jalkanen, A. A. Lindberg*, et al.* (1990). "Salmonella lipopolysaccharide in synovial cells from patients with reactive arthritis." Lancet. **335**(8691): 685-8.

Hermann, E., B. Fleischer, W. J. Mayet*, et al.* (1989). "Response of synovial fluid T cell clones to Yersinia enterocolitica antigens in patients with reactive Yersinia arthritis." Clin Exp Immunol. **75**(3): 365-70.

Hugot, J. P., M. Chamaillard, H. Zouali*, et al.* (2001). "Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease." Nature **411**(6837): 599-603.

Hukuda, S., M. Minami, T. Saito*, et al.* (2001). "Spondyloarthropathies in Japan: nationwide questionnaire survey performed by the Japan Ankylosing Spondylitis Society." J Rheumatol. **28**(3): 554-9.

Kim, T. H., W. S. Uhm and R. D. Inman (2005). "Pathogenesis of ankylosing spondylitis and reactive arthritis." Curr Opin Rheumatol. **17**(4): 400-5.

May, E., E. Marker-Hermann, B. M. Wittig, *et al.* (2000). "Identical T-cell expansions in the colon mucosa and the synovium of a patient with enterogenic spondyloarthropathy." Gastroenterology. **119**(6): 1745-55.

Mielants, H., E. M. Veys, S. Goemaere, *et al.* (1993). "A prospective study of patients with spondyloarthropathy with special reference to HLA-B27 and to gut histology." J Rheumatol. **20**(8): 1353-8.

Mielants, H., E. M. Veys, C. Cuvelier, *et al.* (1995). "The evolution of spondyloarthropathies in relation to gut histology. III. Relation between gut and joint." J Rheumatol **22**(12): 2279-84.

Mielants, H., E. M. Veys, C. Cuvelier, *et al.* (1996). "Course of gut inflammation in spondylarthropathies and therapeutic consequences." Baillieres Clin Rheumatol. **10**(1): 147-64.

Mielants, H., F. De Keyser, D. Baeten, *et al.* (2005). "Gut inflammation in the spondyloarthropathies." Curr Rheumatol Rep. **7**(3): 188-94.

Netea, M. G., G. Ferwerda, D. J. de Jong, *et al.* (2005). "NOD2 3020insC mutation and the pathogenesis of Crohn's disease: impaired IL-1beta production points to a loss-of-function phenotype." Neth J Med. **63**(8): 305-8.

Ogura, Y., D. K. Bonen, N. Inohara, *et al.* (2001). "A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease." Nature **411**(6837): 603-6.

Pacheco-Tena, C., C. Alvarado De La Barrera, Y. Lopez-Vidal, *et al.* (2001). "Bacterial DNA in synovial fluid cells of patients with juvenile onset spondyloarthropathies." Rheumatology (Oxford). **40**(8): 920-7.

Peeters, H., B. Vander Cruyssen, D. Laukens, *et al.* (2004). "Radiological sacroiliitis, a hallmark of spondylitis, is linked with CARD15 gene polymorphisms in patients with Crohn's disease." Ann Rheum Dis. **63**(9): 1131-4.

Reveille, J. D. (2004). "The genetic basis of spondyloarthritis." Curr Rheumatol Rep. **6**(2): 117-25.

Salmi, M. and S. Jalkanen (2001). "Human leukocyte subpopulations from inflamed gut bind to joint vasculature using distinct sets of adhesion molecules." J Immunol. **166**(7): 4650-7.

Taurog, J. D., J. A. Richardson, J. T. Croft, *et al.* (1994). "The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats." J Exp Med. **180**(6): 2359-64.

Ugocsai, P., S. Barlage, A. Dada, *et al.* (2006). "Regulation of surface CD163 expression and cellular effects of receptor mediated hemoglobin-haptoglobin uptake on human

monocytes and macrophages." Cytometry A **14**: 14.

Van Damme, N., F. De Keyser, P. Demetter*, et al.* (2001a). "The proportion of Th1 cells, which prevail in gut mucosa, is decreased in inflammatory bowel syndrome." Clin Exp Immunol. **125**(3): 383-90.

Van Damme, N., M. De Vos, D. Baeten*, et al.* (2001b). "Flow cytometric analysis of gut mucosal lymphocytes supports an impaired Th1 cytokine profile in spondyloarthropathy." Ann Rheum Dis. **60**(5): 495-9.

Van Damme, N., D. Elewaut, D. Baeten*, et al.* (2001c). "Gut mucosal T cell lines from ankylosing spondylitis patients are enriched with alphaEbeta7 integrin." Clin Exp Rheumatol. **19**(6): 681-7.

Van den Bosch, F., E. Kruithof, M. De Vos*, et al.* (2000). "Crohn's disease associated with spondyloarthropathy: effect of TNF-alpha blockade with infliximab on articular symptoms." Lancet. **356**(9244): 1821-2.

Van Den Bosch, F., E. Kruithof, D. Baeten*, et al.* (2002). "Randomized double-blind comparison of chimeric monoclonal antibody to tumor necrosis factor alpha (infliximab) versus placebo in active spondylarthropathy." Arthritis Rheum. **46**(3): 755-65.

Wright, V. (1978). "Seronegative polyarthritis: a unified concept." Arthritis Rheum. **21**(6): 619-33.

Chapter 6

CARD15 GENE POLYMORPHISMS IN PATIENTS WITH SPONDYLOARTHROPATHIES IDENTIFY A SPECIFIC PHENOTYPE PREVIOUSLY RELATED TO CROHN'S DISEASE

Ann Rheum Dis. 2005 Jun; 64(6): 930-5

Debby Laukens[1], Harald Peeters[2], Denis Marichal[2], Bert Vander Cruyssen[3], Herman Mielants[3], Dirk Elewaut[3], Pieter Demetter[4], Claude Cuvelier[4], Marthe Van Den Berghe[3], Pieter Rottiers[1], Eric M Veys[3], Erik Remaut[1], Lothar Steidler[1, 5], Filip De Keyser[3] and Martine De Vos[2]

[1]Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium
[2]Department of Gastroenterology, Ghent University Hospital, Gent, Belgium
[3]Department of Rheumatology, Ghent University Hospital, Gent, Belgium
[4]Department of Pathology, Ghent University Hospital, Gent, Belgium
[5]Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland

ABSTRACT

__Background & aims__*:* Association between spondyloarthropathy (SpA) and Crohn's disease (CD) is a well-known phenomenon. A risk for evolution to CD was already demonstrated in the subgroup of SpA patients with associated chronic gut inflammation. We investigated whether the reported polymorphisms in the *CARD15* gene, a susceptibility gene for CD, are associated with the presence of preclinical intestinal inflammation observed in SpA.
__Methods__*:* We included 104 SpA patients who underwent an ileocolonoscopy with biopsies between 1983 and 2004. Using RFLP-PCR, we assessed the prevalence of three single nucleotide polymorphisms in the *CARD15* gene (R702W, G908R and 1007fs) and compared them to an ethnically matched CD population and a control population.
__Results__*:* The carrier frequency of R702W, G908R or 1007fs variants in the SpA populations (20%) was similar as in the control population (17%), but increased to 38% in the subgroup of

SpA patients with chronic gut inflammation. This was significantly higher than in the other SpA subgroups (*P*=0.001) and the control group (*P*=0.006) but not significantly different from the prevalence in CD (49%). This indicates that *CARD15* polymorphisms are associated with a higher risk for development of chronic gut inflammation.

Conclusions: *CARD15* gene polymorphisms clearly identify a subgroup of patients with SpA associated with chronic intestinal inflammation.

INTRODUCTION

The spondyloarthropathies are a group of interrelated inflammatory diseases characterized by a pauciarticular, peripheral, asymmetrical arthritis and/or axial involvement with ankylosing spondylitis (AS) as prototype (Lawrence *et al.* 1998; Khan 2002). Reported prevalences of spondyloarthropathies vary between 0.2 and 1.9% (Braun *et al.* 1998). Although association with *HLA-B27* is strong, recent genetic studies suggest a polygenic model of susceptibility (Brown *et al.* 1997; Said-Nahal *et al.* 2000; Reveille *et al.* 2001; Granfors *et al.* 2002).

In up to 60% of spondyloarthropathy (SpA) patients, articular involvement is associated with subclinical histological evidence of chronic or acute gut inflammation in ileum or colon (De Vos 1989; Simenon *et al.* 1990; Leirisalo-Repo *et al.* 1994). We described a long-term evolution to overt Crohn's disease (CD) in 13% of patients with initial chronic gut inflammation (Mielants *et al.* 1995a; De Vos *et al.* 1996). The presence of chronic intestinal inflammation did not relate to *HLA-B27*, but a weak association was found with *HLA-B62* (Mielants *et al.* 1995a).

The observed immunologic similarities in SpA with gut inflammation and CD support the concept that this subgroup of SpA patients can be considered as a model for early immune alterations related to CD. An enrichment of gut mucosal T cell lines with αEβ7 integrin and an increased expression of its ligand, E-cadherin, was found in intestine of CD as well as SpA patients (Elewaut *et al.* 1998a; Demetter *et al.* 2000; Demetter *et al.* 2002). Re-circulation of T cells primed in the gut to synovial tissue is one potential mechanism by which gut and synovial inflammation could be linked. This hypothesis is supported by an altered expression of β7 integrins, which are highly expressed within the gut, on synovial T cells from SpA patients compared to rheumatoid arthritis (Elewaut *et al.* 1998b). Another potential mechanism includes trafficking of antigen presenting cells between gut and joint. Consistent with this was the augmented infiltration of gut mucosa and synovium with $CD163^{+}$ macrophages (producing IL1 and TNF) in both CD and SpA patients (Baeten *et al.* 2002; Demetter *et al.* 2005). Finally, a comparable beneficial clinical effect of infliximab, a monoclonal antibody to TNF suggests a

key role of this cytokine in both diseases (Van den Bosch *et al.* 2000; Van Den Bosch F 2002).

In 2001, a correlation was reported between polymorphisms in the *CARD15* gene and an increased susceptibility for CD (Hampe *et al.* 2001; Hugot *et al.* 2001; Ogura *et al.* 2001a). Three independent single nucleotide polymorphisms (SNPs) in *CARD15* are associated with CD in about 30 to 46% of patients (1 frameshift mutation, 1007fs (SNP13), and 2 missense mutations, R702W (SNP8) and G908R (SNP12)) (Hugot *et al.* 2001; Esters *et al.* 2004). These variants increase the risk for CD by a factor of 3 for heterozygous and by a factor of 38 or 44 for respectively homozygous or compound heterozygous individuals (Hugot *et al.* 2001). Lower prevalences have been described in CD patients in Scotland, Ireland and Northern Europe, whereas no association could be found in Japan (Inoue *et al.* 2002; Yamazaki *et al.* 2002; Helio *et al.* 2003; Arnott *et al.* 2004).

CARD15 encodes for an intracellular protein, which is expressed in monocytes, granulocytes, dendritic, epithelial and paneth cells, and has binding affinity for bacterial cell wall components like muramyldipeptides (Girardin *et al.* 2003). The *CARD15* protein is involved in NFκB activation and in apoptosis by two N-terminal Caspase Recruitment Domains (hence the term CARD), although its precise pathogenic role in CD remains to be determined (Ogura *et al.* 2001b; Bonen *et al.* 2003; Girardin *et al.* 2003).

CARD15 gene polymorphisms have also been linked with another related syndrome, Blau's syndrome, characterized by granulomatous inflammation of uvea, skin and joints (Miceli-Richard *et al.* 2001).

Several studies have been performed to investigate the role of *CARD15* polymorphisms in SpA. These studies did not demonstrate an association with SpA or AS in particular (D'Amato 2002; Miceli-Richard *et al.* 2002; Breban *et al.* 2003; Ferreiros-Vidal *et al.* 2003; van der Paardt *et al.* 2003). Yet an increased prevalence of *CARD15* polymorphisms was found in psoriatic arthritis but not in psoriatic skin disease (Borgiani *et al.* 2002; Rahman *et al.* 2003; Young *et al.* 2003). A recent Italian study however could not confirm this association (Giardina *et al.* 2004). Nevertheless, this finding could emphasize the importance of investigating the possible role of these genetic variants in specific, clinical subpopulations of patients. In CD as well, *CARD15* polymorphisms seem to be related with certain clinical phenotypes (Abreu *et al.* 2002; Ahmad *et al.* 2002; Cuthbert *et al.* 2002; Lesage *et al.* 2002; Peeters *et al.* 2004).

In view of the apparent correlation between gut inflammation in SpA and clinical evolution to CD, we investigated whether the presence of polymorphisms in this susceptibility gene for CD would be associated with gut inflammation in SpA patients.

MATERIALS AND METHODS

Study population

This study included 104 Caucasian SpA patients (according to the ESSG criteria (Dougados *et al.* 1991), who underwent an ileocolonoscopy with concomitant ileal and colonic biopsies between 1983 and 2004. This population consisted of 74 male and 30 female patients with a mean age of 46 years (range: 21-77 yrs). SpA patients were systematically referred by the rheumatologist for an ileocolonoscopy with biopsies, independent of the presence of GI symptoms.

Patients with the diagnosis of clinical Crohn's disease or psoriasis prior to the diagnosis of SpA, were excluded from the study.

A subgroup of 54 patients, all having a long-term follow-up since their diagnosis of SpA (ranging from 17 to 49 years), was recently clinically reassessed. New follow-up colonoscopies were not performed.

The total SpA population consisted of 75 patients with ankylosing spondylitis (AS) according to the modified New York criteria (van der Linden *et al.* 1984) and 29 patients with an undifferentiated form of SpA (uSpA). Eighteen AS patients only had axial involvement, whereas 57 AS patients also had peripheral disease (defined as the history or presence of peripheral arthritis and/or enthesitis). Twenty-five uSpA patients had peripheral disease and 4 uSpA patients only had axial involvement. These 4 patients had inflammatory low back pain and fulfilled the ESSG criteria, however not the modified New York criteria for AS.

HLA-B27 status was known in a total of 81 patients. In 53 patients both *HLA-B27* and *HLA-B62* status were known.

A population of 156 consecutive patients with proven CD on clinical, endoscopical and histological grounds was included as well. This cohort included 57 male and 99 female patients with a mean age of 38 years (range: 18-80 yrs).

Prevalences were also compared to those observed in a control population including 140 individuals.

The study was approved by the local ethics committee. All patients signed an informed consent.

Histological classification

A classification of histologic lesions was used as reported in previous studies (De Vos 1989; Mielants *et al.* 1995a; Mielants *et al.* 1995b; Mielants *et al.* 1995c; De Vos *et al.* 1996). Three subgroups were distinguished: patients with normal gut histology, acute and chronic

inflammation (Cuvelier *et al.* 1987). In acute inflammatory lesions normal architecture was well preserved. A mucosal and epithelial infiltration by neutrophils and eosinophils was found, without a significant increase in lymphocytes. Small superficial ulcers covered with fibrin and neutrophils overlying hyperplastic lymphoid follicles were occasionally observed. The lamina propria was oedematous and hemorrhagic and contained mainly polymorphonuclear cells. The pattern of inflammation was similar to that seen in acute self-limiting bacterial enterocolitis.

The principal features of chronic inflammatory lesions were mucosal architectural alterations with crypt distortion and atrophy in the colon and villous blunting and fusion in ileal mucosa. Both in ileum and colon there was an increased mixed cellularity and formation of basal lymphoid aggregates in the lamina propria.

Whenever one of several biopsies featured chronic lesions, regardless of acute or active inflammation in other fragments, a diagnosis of chronic inflammation was made.

Although NSAID may induce intestinal disorders, we and others excluded these drugs as aetiology of reported chronic inflammation (De Vos 1989; Simenon *et al.* 1990; Altomonte *et al.* 1994).

CARD15 genotyping (R702W, G908R and 1007fs), _HLA-B27_ and _HLA-B62_ typing

Genomic DNA was extracted from whole blood using Qiagen blood and cell culture DNA kit (Westburg BV, Leusden, The Netherlands) and genotyped all patients for R702W, G908R and 1007fs using RFLP-PCR, followed by separation of the DNA fragments on a 2.5% agarose gel. The missense mutation R702W (GenBank accession number G67950) abolishes the restriction site for *Msp*I, resulting in an intact 130-bp band for mutant alleles compared to two bands of 54- and 76-bp for wild type alleles (forward primer: 5'-CAG CCC TGA TGA CAT TTC TCT T-3', reverse primer: 5'-AGC CGC TCC TCC TGC ATC TCG TA-3'). The missense mutation G908R (GenBank accession number G67951) creates a restriction site for *HinP*1l. The frameshift mutation 1007fs (GenBank accession number G67955) creates a restriction site for *Nla*IV. The presence of a mutant allele results in two bands of 219 and 41 bp, while the wild type allele produces a single 260-bp product (forward primer: 5'-CTG AGC CTT TGT TGA TGA GC-3', reverse primer: 5'-TCT TCC AAC CAC ATC CCC ATT-3').

In the patients with a known *HLA-B27* and *HLA-B62* status, typing of these markers had been performed using the micro-lymphocytotoxity test according to Terasaki and McClelland (Cuthbert *et al.* 2002).

Statistical analysis

Statistical significance was determined by the χ^2 test and Odds Ratio using SPSS (SPSS inc., Chicago, Illinois). Multivariate analysis (logistic regression) was performed to investigate whether an association, found through univariate analysis, was independent from other genetic markers. *P* values less than 0.05 were considered significant.

RESULTS

We subdivided our cohort in three groups according to the gut histology. Forty patients (38%) had a normal histology, 24 patients (23%) had acute gut inflammation and 40 (38%) showed chronic gut inflammation (Table 1).

Table 1 Prevalence of *CARD15* variants in the populations, according to subtypes defined at baseline

	N	Carriers of *CARD15* variant(s)
Classification		
control	140	24 (17%)
crohn	156	77 (49%)*
spondyloarthropathy (SpA)	104	21 (20%)§
ankylosing Spondylitis (AS)	75	16 (21%)
undifferentiated SpA (uSpA)	29	5 (17%)
Gut histology in SpA population		
chronic inflammation	40 (38%)	15 (38%)‡
acute inflammation	24 (23%)	0 (0%)
normal histology	40 (38%)	6 (15%)

*Chi-square: *P<0.001* (carrier frequency in CD vs control population)
§Chi-square: *P*=0.5 (carrier frequency in general SpA vs control population)
‡Chi-square: *P*=0.001(chronic inflammation in patients with *CARD15* variant vs. chronic inflammation in those without *CARD15* polymorphisms

Univariate analysis

Prevalence of CARD15 polymorphisms in the SpA, CD and control populations. The prevalences of *CARD15* polymorphisms in the total SpA (20%), specific AS (21%) and uSpA (17%) population did not differ significantly (Table 1). All except one (homozygous for the 1007fs allelic variant) were heterozygous for at least one mutation. The prevalence of R702W, G908R and 1007fs allelic variants in this SpA population was 12%, 4% and 5% respectively (Table 2).
No compound heterozygousity was found. All carriers of *CARD15* polymorphisms in the SpA group had (a history of) peripheral disease (Table 3). There were no significant differences

Table 2 Carrier frequency of *CARD15* variants in patients with SpA, CD and controls (%)

	SpA (*N*=104)				CD (*N*=156)				Controls (*N*=140)			
	R702W	G908R	1007fs	overall*	R702W	G908R	1007fs	overall*	R702W	G908R	1007fs	overall*
CARD15$^{-/+}$	12 (12)	4 (4)	4 (4)	21 (20)	39 (25)	11 (7)	27 (17)	77 (49)	18 (13)	1 (1)	6 (4)	24 (17)
CARD15$^{-/-}$	0 (0)	0 (0)	1 (1)		4 (3)	3 (2)	0 (0)		0 (0)	0 (0)	0 (0)	

Number of patients carrying R702W, G908R or 1007fs variants
CARD15-/+: heterozygous; *CARD15*-/-: homozygous
Overall = total number of patients in the group carrying at least 1 variant
* The sum of all allelic *CARD15* variants is greater than the overall number of patients at least carrying one variant, since some patients carry 2 different SNP's, thus displaying a compound heterozygous status

concerning the disease duration and the duration of the follow-up period between the SpA patients carrying *CARD15* polymorphisms and the group of patients without these polymorphisms (data not shown).

In the CD population, a carrier frequency of 49% (77 of 156 patients) was observed (Table 1). Forty-three CD patients carried at least one R702W polymorphism, 14 patients carried at least one G908R polymorphism and 27 patients carried at least one 1007fs polymorphism. Fourteen patients carried two polymorphisms of which 7 patients were homozygous and 7 patients compound heterozygous (Table 2).

In the control group, 24 individuals (17%) carried *CARD15* polymorphisms (Table 1). All except one (compound heterozygous for the R702W and 1007fs variant) were single heterozygous (Table 2).

The prevalence of polymorphisms in the SpA cohort (20%) was not different from that observed in the control group (17%) (OR: 1.22, CI: 0.64–2.34, P=0.5) and significantly lower compared to the prevalence found in our CD population (49%) (OR: 3.85, CI: 2.17-6.83, P<0.001).

Association between CARD15 polymorphisms and intestinal inflammation in SpA patients. The carrier frequency in the subpopulation of SpA patients with chronic gut inflammation was 38% (15 of 40 patients) which was significantly higher compared to the control population (OR: 2.9, CI: 1.33-6.30, P=0.006) and the other SpA populations (OR: 5.80, CI: 2.02-16.68, P=0.001) and not statistically different from that observed in our CD population (49%, OR: 1.62, CI: 0.80-3.31, P=0.2) (Table 1).

Of all SpA patients carrying *CARD15* polymorphisms, 71% (15 out of 21 patients) had chronic gut inflammation, 0% acute inflammation and 29% presented with normal histology (Table 1). The only SpA patient carrying two *CARD15* variants also had chronic gut inflammation. In

contrast, only 25 out of 83 patients with a wild type genotype (30%) had chronic gut inflammation, 29% acute inflammation and 41% normal histology. Consequently, the presence of *CARD15* polymorphisms was associated with a higher risk for development of chronic gut inflammation.

There are no statistically significant differences between the AS and the uSpA group concerning the prevalence of *CARD15* polymorphisms in patients with normal (3/29 in AS vs 3/11 in uSpA, OR: 3.3, CI: 0.5-19.4, *P*=0.3), acute (0/13 in AS vs 0/11 in uSpA) or chronic (13/33 in AS vs 2/7 in uSpA, OR: 1.6, CI: 0.3-9.7, *P*=0.7) gut inflammation.

In the subgroup of 54 patients who were clinically reassessed, 4 patients evolved from histological chronic gut inflammation towards clinically overt Crohn's disease. Two of these 4 patients carried *CARD15* polymorphisms. The other 22 patients with chronic gut inflammation in this group did not develop clinical CD.

Association between CARD15 polymorphisms and HLA-B27 in SpA patients. There was no significant association between the presence of these 2 genetic markers. Six of 34 *HLA-B27* negative patients carried *CARD15* polymorphisms versus 13 of 47 *HLA-B27* positive patients (OR: 1.8, CI: 0.6-5.3, *P*=0.3).

Multivariate analysis

In the subgroup of 53 SpA patients of whom both *HLA-B27* and *HLA-B62* status were known, logistic regression was performed (with the presence of chronic gut inflammation as dependent variable). This showed that the association between chronic gut inflammation and *CARD15* polymorphisms (OR: 17.3, CI: 2.0-152.3, *P*=0.01) is independent of *HLA-B27* (OR: 1.7, CI: 0.5-6.0, *P*=0.42) and *HLA-B62* (OR: 2.5, CI: 0.5-13.0, *P*=0.28).

CONCLUSION

This study describes a novel and remarkably strong association between variants in a host defence gene located on chromosome 16, *CARD15*, and a chronic form of gut inflammation in patients with SpA. Interestingly, the prevalence of *CARD15* polymorphisms in this subgroup of SpA patients was not significantly different from that observed in patients suffering from CD.

Three single nucleotide polymorphisms have been associated with CD (Hampe *et al.* 2001; Hugot *et al.* 2001; Ogura *et al.* 2001a). One variant (1007fs) encodes a truncated protein which results in altered activation of NFκB in response to bacterial stimuli (Ogura *et al.* 2001b; Bonen

et al. 2003; Girardin *et al.* 2003). The two other single nucleotide polymorphisms (R702W and G908R) result in an amino acid substitution.

Table 3 Prevalences of *CARD15* polymorphisms according to the presence of mainly axial or peripheral involvement in the total SpA group, AS group and uSpA group

total SpA group (*N*=104)			
	CARD15		*P*=0.006
	wild type	variant	total
axial	22	0	22
peripheral	61	21	82
total	83	21	104

AS group (*N*=75)			
	CARD15		*P*=0.006
	wild type	variant	total
axial	18	0	18
peripheral	41	16	57
total	59	16	75

uSpA group (*N*=29)			
	CARD15		*P*=1.0
	wild type	variant	total
axial	4	0	4
peripheral	20	5	25
total	24	5	29

More recently, several groups assessed the linkage of *CARD15* variants in CD to particular clinical phenotypes but the results of these retrospective studies are disparate. The presence of two mutations has been linked to younger age at onset and preferential involvement of small bowel (Lesage *et al.* 2002). Preference for ileal involvement was also reported by Cuthbert (Cuthbert *et al.* 2002) and by Ahmad (Ahmad *et al.* 2002). Prevalence for fibrostenosing disease was dominant in a study of Abreu (Abreu *et al.* 2002). In these studies, no association of *CARD15* variants with extra-intestinal involvement could be retained.

The present study demonstrates a new association between these three CD-associated variants in the leucine-rich regulatory domain of the *CARD15* gene and a distinct subpopulation of patients with spondyloarthropathies. Similarly to previous reports, the overall prevalence of mutations in SpA patients was not statistically different from the prevalence in our control population (D'Amato 2002; Miceli-Richard *et al.* 2002; Breban *et al.* 2003; Ferreiros-Vidal *et al.* 2003; van der Paardt *et al.* 2003). However, unlike the previous studies, we identified a distinct clinical subgroup, characterised by the presence of chronic inflammatory gut lesions, with a remarkably high prevalence (38%) of *CARD15* polymorphisms, being not significantly different from the prevalence in the CD population (48%) and significantly higher compared to the control population (17%) and the other SpA patients. Previous studies from our group showed that in particular these patients with chronic gut inflammation were at risk for the progression to CD (Mielants *et al.* 1995a; De Vos *et al.* 1996).

Striking in the present study, none of the SpA patients with only axial disease carried *CARD15* polymorphisms. Carriers of these polymorphisms all had (a history of) peripheral disease. This is in concordance with previous studies, where more chronic gut inflammation could be found in AS patients with peripheral disease compared to strict axial AS patients (Mielants *et al.* 1995a).

One previous study investigated *CARD15* polymorphisms in AS patients with CD and ulcerative colitis (UC) (Crane *et al.* 2002). It did not show a higher prevalence of *CARD15* variants in AS patients with CD compared to idiopathic AS, AS with UC or healthy controls. However, the low prevalence of *CARD15* variants in the CD population with AS was not compared to the prevalence in a general CD population and it unexpectedly revealed a possible association between the G908R *CARD15* variant and AS patients with ulcerative colitis.

Moreover, in a recent study we found an association between *CARD15* polymorphisms and the presence of radiological sacroiliitis in CD patients, unrelated to the *HLA-B27* status of these subjects. These data already pointed at a role for the *CARD15* gene in the link between gut and joint inflammation (Peeters *et al.* 2004).

Our findings confirm the previous reported clinical, therapeutical and immunological links between SpA and CD and provide also genetic proof for the association between both diseases. Since the chronic gut inflammation in the majority of SpA patients remains asymptomatic, this might suggest that *CARD15* polymorphisms could be linked with the development of (subclinical) chronic gut inflammation rather than with CD as such.

The underlying pathogenetic mechanisms that could explain the phenotypic expression of *CARD15* mutations in SpA need to be investigated. *CARD15* encodes a cytosolic protein that could play a role in SpA by interference with transport of antigens by macrophages from mucosal surfaces to the joints (Salmi *et al.* 2001). *CARD15* seems to function as an intracellular receptor for bacterial components, where the C-terminal leucine-rich regulatory domain is crucial for responsiveness. Cellular response to bacterial products was altered in HEK293T cells transfected with expression plasmids containing any of the three SNPs (Ogura *et al.* 2001b; Bonen *et al.* 2003; Girardin *et al.* 2003). Moreover, expression of *CARD15* in myeloblastic and epithelial cells is enhanced by pro-inflammatory cytokines and bacterial components via NFκB (Ogura *et al.* 2001b; Gutierrez *et al.* 2002; Bonen *et al.* 2003). This response is likely to mediate cytokine production including TNF suggesting that up-regulation of CARD15 may be part of a positive regulatory loop and facilitate the response of the host to pathogens. A genetically determined disturbed handling of bacterial products in the intestinal tract, leading to an altered transport of antigens by T cells to synovial tissue, is an interesting hypothesis that should be investigated in spondyloarthropathy. A further identification and characterisation of inflammatory cells involved in gut and joint inflammation may also lead to new therapeutic targets.

In conclusion, a distinct phenotype associated with the three main CD associated *CARD15* variants is reported in patients with SpA. Our data show that the presence of *CARD15* variants in

SpA patients strongly predisposes to chronic intestinal inflammation, defining a population at risk for evolution to CD. However, the persistence of the subclinical character of the inflammation in a large part of patients may reflect that CD is a multigenic disease or alternatively that the heterozygous carriage of *CARD15* polymorphisms predisposes only to a subclinical inflammation.

REFERENCES

Abreu, M. T., K. D. Taylor, Y. C. Lin*, et al.* (2002). "Mutations in NOD2 are associated with fibrostenosing disease in patients with Crohn's disease." Gastroenterology **123**(3): 679-88.

Ahmad, T., A. Armuzzi, M. Bunce*, et al.* (2002). "The molecular classification of the clinical manifestations of Crohn's disease." Gastroenterology **122**(4): 854-66.

Altomonte, L., A. Zoli, A. Veneziani*, et al.* (1994). "Clinically silent inflammatory gut lesions in undifferentiated spondyloarthropathies." Clin Rheumatol. **13**(4): 565-70.

Arnott, I. D., E. R. Nimmo, H. E. Drummond*, et al.* (2004). "NOD2/CARD15, TLR4 and CD14 mutations in Scottish and Irish Crohn's disease patients: evidence for genetic heterogeneity within Europe?" Genes Immun. **5**(5): 417-25.

Baeten, D., P. Demetter, C. A. Cuvelier*, et al.* (2002). "Macrophages expressing the scavenger receptor CD163: a link between immune alterations of the gut and synovial inflammation in spondyloarthropathy." J Pathol **196**(3): 343-50.

Bonen, D. K., Y. Ogura, D. L. Nicolae*, et al.* (2003). "Crohn's disease-associated NOD2 variants share a signaling defect in response to lipopolysaccharide and peptidoglycan." Gastroenterology **124**(1): 140-6.

Borgiani, P., L. Vallo, M. R. D'Apice*, et al.* (2002). "Exclusion of CARD15/NOD2 as a candidate susceptibility gene to psoriasis in the Italian population." Eur J Dermatol **12**(6): 540-2.

Braun, J., M. Bollow, G. Remlinger*, et al.* (1998). "Prevalence of spondylarthropathies in HLA-B27 positive and negative blood donors." Arthritis Rheum **41**(1): 58-67.

Breban, M., R. Said-Nahal, J. P. Hugot*, et al.* (2003). "Familial and genetic aspects of spondyloarthropathy." Rheum Dis Clin North Am **29**(3): 575-94.

Brown, M. A., L. G. Kennedy, A. J. MacGregor*, et al.* (1997). "Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment." Arthritis Rheum **40**(10): 1823-8.

Crane, A. M., L. Bradbury, D. A. van Heel*, et al.* (2002). "Role of NOD2 variants in spondylarthritis." Arthritis Rheum **46**(6): 1629-33.

Cuthbert, A. P., S. A. Fisher, M. M. Mirza*, et al.* (2002). "The contribution of NOD2 gene

mutations to the risk and site of disease in inflammatory bowel disease." Gastroenterology **122**(4): 867-74.

Cuvelier, C., C. Barbatis, H. Mielants*, et al.* (1987). "Histopathology of intestinal inflammation related to reactive arthritis." Gut **28**(4): 394-401.

D'Amato, M. (2002). "The Crohn's associated NOD2 3020InsC frameshift mutation does not confer susceptibility to ankylosing spondylitis." J Rheumatol **29**(11): 2470-1.

De Vos, M. (1989). "Ileocolonoscopy in seronegative spondylarthropathy." Gastroenterology **96**: 339-344.

De Vos, M., H. Mielants, C. Cuvelier*, et al.* (1996). "Long-term evolution of gut inflammation in patients with spondyloarthropathy." Gastroenterology **110**(6): 1696-703.

Demetter, P., D. Baeten, F. De Keyser*, et al.* (2000). "Subclinical gut inflammation in spondyloarthropathy patients is associated with upregulation of the E-cadherin/catenin complex." Ann Rheum Dis **59**(3): 211-6.

Demetter, P., J. A. Van Huysse, F. De Keyser*, et al.* (2002). "Increase in lymphoid follicles and leukocyte adhesion molecules emphasizes a role for the gut in spondyloarthropathy pathogenesis." J Pathol **198**(4): 517-22.

Demetter, P., M. De Vos, J. A. Van Huysse*, et al.* (2005). "Colon mucosa of patients both with spondyloarthritis and Crohn's disease is enriched with macrophages expressing the scavenger receptor CD163." Ann Rheum Dis. **64**(2): 321-4.

Dougados, M., S. van der Linden, R. Juhlin*, et al.* (1991). "The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy." Arthritis Rheum **34**(10): 1218-27.

Elewaut, D., F. De Keyser, C. Cuvelier*, et al.* (1998a). "Distinctive activated cellular subsets in colon from patients with Crohn's disease and ulcerative colitis." Scand J Gastroenterol **33**(7): 743-8.

Elewaut, D., F. De Keyser, F. Van Den Bosch*, et al.* (1998b). "Enrichment of T cells carrying beta7 integrins in inflamed synovial tissue from patients with early spondyloarthropathy, compared to rheumatoid arthritis." J Rheumatol **25**(10): 1932-7.

Esters, N., M. Pierik, K. van Steen*, et al.* (2004). "Transmission of CARD15 (NOD2) variants within families of patients with inflammatory bowel disease." Am J Gastroenterol **99**(2): 299-305.

Ferreiros-Vidal, I., J. Amarelo, F. Barros*, et al.* (2003). "Lack of Association of Ankylosing Spondylitis with the Most Common NOD2 Susceptibility Alleles to Crohn's Disease." J Rheumatol **30**(1): 102-4.

Giardina, E., G. Novelli, A. Costanzo*, et al.* (2004). "Psoriatic arthritis and CARD15 gene polymorphisms: no evidence for association in the Italian population." J Invest Dermatol. **122**(5): 1106-7.

Girardin, S. E., I. G. Boneca, J. Viala*, et al.* (2003). "Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection." J Biol Chem **278**(11): 8869-72.

Granfors, K., E. Marker-Hermann, F. de Keyser*, et al.* (2002). "The cutting edge of spondylarthropathy research in the millennium." Arthritis Rheum **46**(3): 606-13.

Gutierrez, O., C. Pipaon, N. Inohara*, et al.* (2002). "Induction of Nod2 in myelomonocytic and intestinal epithelial cells via nuclear factor-kappa B activation." J Biol Chem **277**(44): 41701-5.

Hampe, J., A. Cuthbert, P. J. Croucher*, et al.* (2001). "Association between insertion mutation in NOD2 gene and Crohn's disease in German and British populations." Lancet **357**(9272): 1925-8.

Helio, T., L. Halme, M. Lappalainen*, et al.* (2003). "CARD15/NOD2 gene variants are associated with familially occurring and complicated forms of Crohn's disease." Gut **52**(4): 558-62.

Hugot, J. P., M. Chamaillard, H. Zouali*, et al.* (2001). "Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease." Nature **411**(6837): 599-603.

Inoue, N., K. Tamura, Y. Kinouchi*, et al.* (2002). "Lack of common NOD2 variants in Japanese patients with Crohn's disease." Gastroenterology **123**(1): 86-91.

Khan, M. A. (2002). "Update on spondyloarthropathies." Ann Intern Med **136**(12): 896-907.

Lawrence, R. C., C. G. Helmick, F. C. Arnett*, et al.* (1998). "Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States." Arthritis Rheum. **41**(5): 778-99.

Leirisalo-Repo, M., U. Turunen, S. Stenman*, et al.* (1994). "High frequency of silent inflammatory bowel disease in spondylarthropathy." Arthritis Rheum **37**(1): 23-31.

Lesage, S., H. Zouali, J. P. Cezard*, et al.* (2002). "CARD15/NOD2 mutational analysis and genotype-phenotype correlation in 612 patients with inflammatory bowel disease." Am J Hum Genet **70**(4): 845-57.

Miceli-Richard, C., S. Lesage, M. Rybojad*, et al.* (2001). "CARD15 mutations in Blau syndrome." Nat Genet **29**(1): 19-20.

Miceli-Richard, C., H. Zouali, S. Lesage*, et al.* (2002). "CARD15/NOD2 analyses in spondylarthropathy." Arthritis Rheum **46**(5): 1405-6.

Mielants, H., E. M. Veys, C. Cuvelier*, et al.* (1995a). "The evolution of spondyloarthropathies in relation to gut histology. II. Histological aspects." J Rheumatol **22**(12): 2273-8.

Mielants, H., E. M. Veys, C. Cuvelier*, et al.* (1995b). "The evolution of spondyloarthropathies in relation to gut histology. III. Relation between gut and joint." J Rheumatol **22**(12): 2279-84.

Mielants, H., E. M. Veys, M. De Vos*, et al.* (1995c). "The evolution of spondyloarthropathies in relation to gut histology. I. Clinical aspects." J Rheumatol **22**(12): 2266-72.

Ogura, Y., D. K. Bonen, N. Inohara*, et al.* (2001a). "A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease." Nature **411**(6837): 603-6.

Ogura, Y., N. Inohara, A. Benito*, et al.* (2001b). "Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB." J Biol Chem **276**(7): 4812-8.

Peeters, H., B. Vander Cruyssen, D. Laukens*, et al.* (2004). "Radiological sacroiliitis, a hallmark of spondylitis, is linked with CARD15 gene polymorphisms in patients with Crohn's disease." Ann Rheum Dis **63**(9): 1131-4.

Rahman, P., S. Bartlett, F. Siannis*, et al.* (2003). "CARD15: a pleiotropic autoimmune gene that confers susceptibility to psoriatic arthritis." Am J Hum Genet **73**(3): 677-81.

Reveille, J. D., E. J. Ball and M. A. Khan (2001). "HLA-B27 and genetic predisposing factors in spondyloarthropathies." Curr Opin Rheumatol **13**(4): 265-72.

Said-Nahal, R., C. Miceli-Richard, J. M. Berthelot*, et al.* (2000). "The familial form of spondylarthropathy: a clinical study of 115 multiplex families. Groupe Francais d'Etude Genetique des Spondylarthropathies." Arthritis Rheum **43**(6): 1356-65.

Salmi, M. and S. Jalkanen (2001). "Human leukocyte subpopulations from inflamed gut bind to joint vasculature using distinct sets of adhesion molecules." J Immunol **166**(7): 4650-7.

Simenon, G., A. Van Gossum, M. Adler*, et al.* (1990). "Macroscopic and microscopic gut lesions in seronegative spondyloarthropathies." J Rheumatol **17**(11): 1491-4.

Van den Bosch, F., E. Kruithof, M. De Vos*, et al.* (2000). "Crohn's disease associated with spondyloarthropathy: effect of TNF-alpha blockade with infliximab on articular symptoms." Lancet. **356**(9244): 1821-2.

Van Den Bosch F, K. E., Baeten D, Herssens A, de Keyser F, Mielants H, Veys EM. (2002). "Randomized double-blind comparison of chimeric monoclonal antibody to tumor necrosis factor alpha (infliximab) versus placebo in active spondylarthropathy." Arthritis Rheum **46**(3): 755-65.

van der Linden, S. M. and M. A. Khan (1984). "The risk of ankylosing spondylitis in HLA-B27 positive individuals: a reappraisal." J Rheumatol **11**(6): 727-8.

van der Paardt, M., J. B. Crusius, M. H. de Koning*, et al.* (2003). "CARD15 gene mutations are not associated with ankylosing spondylitis." Genes Immun **4**(1): 77-8.

Yamazaki, K., M. Takazoe, T. Tanaka*, et al.* (2002). "Absence of mutation in the NOD2/CARD15 gene among 483 Japanese patients with Crohn's disease." J Hum Genet **47**(9): 469-72.

Young, C., M. H. Allen, A. Cuthbert*, et al.* (2003). "A Crohn's disease-associated insertion polymorphism (3020insC) in the NOD2 gene is not associated with psoriasis vulgaris, palmo-plantar pustular psoriasis or guttate psoriasis." Exp Dermatol **12**(4): 506-9.

Chapter 7

ALTERED GUT TRANSCRIPTOME IN SPONDYLOARTHROPATHY

Ann Rheum Dis. 2006 Feb

Debby Laukens[1], Harald Peeters[2], Bert Vander Cruyssen[3], Tom Boonefaes[1], Dirk Elewaut[3], Filip De Keyser[3], Herman Mielants[3], Claude Cuvelier[4], Eric M. Veys[3], Kenny Knecht[5], Paul Van Hummelen[6], Erik Remaut[1], Lothar Steidler[1, 7], Martine De Vos[2] and Pieter Rottiers[1]

[1]Department for Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium
[2]Department of Gastroenterology, Ghent University Hospital, Gent, Belgium
[3]Department of Rheumatology, Ghent University Hospital, Gent, Belgium
[4]Department of Pathology, Ghent University Hospital, Gent, Belgium
[5]Applied Maths BVBA, Sint-Martens-Latem, Belgium
[6]MicroArray Facility, VIB, Leuven, Belgium
[7]Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland

ABSTRACT

Background & aims: Intestinal inflammation is a common feature of spondyloarthropathy (SpA) and Crohn's disease (CD). Inflammation is manifested clinically in CD, and subclinical in SpA. However, a fraction of SpA patients develop overt CD. The aim of this study was to investigate whether subclinical gut lesions in SpA patients are associated with transcriptome changes comparable to those seen in CD. We examined global gene expression in non-inflamed colon biopsies, and screened for differentially expressed genes.

Methods: Macroarray analysis was used as an initial genome-wide screen for selecting a comprehensive set of genes relevant to CD and SpA. This led to the identification of 2,625 expressed sequence tags (ESTs) that are differentially expressed in the colon of CD and/or SpA patients. These clones, together with appropriate controls (6,779 in total) were used to construct a glass-based microarray, which was then used to analyze colon biopsies from 15 SpA patients,

11 CD patients, and 10 controls.

Results: Ninety-five genes were identified as differentially expressed in SpA patients with a history of subclinical chronic gut inflammation as well as in CD patients. Principal component analysis of this filtered set of genes successfully distinguished colon biopsies from the three groups studied. Spondyloarthropathy patients with subclinical chronic gut inflammation cluster together, and are more related to CD.

Conclusions: The transcriptome in the intestine of SpA patients differs from that of controls. Moreover, these gene alterations are comparable to those seen in CD, confirming initial clinical observations. Based on these findings, new (genetic) markers for detection of early CD in SpA patients can be considered.

INTRODUCTION

The clinical association between Spondyloarthropathy (SpA) and Crohn's disease (CD) is illustrated by the concurrence of similar arthropathy and intestinal inflammation in the two diseases, indicating a shared etiology and pathogenesis. Depending on the imaging technique used, up to one third of CD patients have peripheral and/or sacroiliac joint abnormalities similar to those seen in various SpA subgroups (Davis *et al.* 1978; Scott *et al.* 1990). In addition, 60% of SpA patients who have no evidence of CD exhibit endoscopic and/or histological signs of subclinical gut inflammation (Mielants *et al.* 1995a). In general, two types of inflammation are observed: acute inflammation as seen in infectious colitis, and chronic inflammation resembling that in CD (Mielants *et al.* 1995a). A striking parallel exists between the activity of inflammation at the joints and the intestine. Moreover, long-term evolution to CD was observed in 13% of SpA patients with initial chronic gut inflammation, supporting the concept of preclinical CD in those patients (De Vos *et al.* 1996). Since these clinical observations, several studies provided additional evidence for a joint-gut axis on the molecular as well as the genetic level. The early immune alterations observed are up-regulation of $\alpha E\beta 7$ integrin on T-cell lines from SpA patients (Elewaut *et al.* 1998b), and an increase in lymphoid follicles and lamina propria mononuclear cells in intestinal biopsies (Demetter *et al.* 2000; Demetter *et al.* 2002). Increased expression of $\alpha E\beta 7$ and the E-cadherin/catenin complex was found in gut mucosa from CD and SpA patients (Elewaut *et al.* 1998a; Demetter *et al.* 2000). A specific subset of $CD163^+$ macrophages is augmented in both groups of patients, supporting the hypothesis of a recirculation of similar clones in the intestinal mucosa and synovium (Baeten *et al.* 2002).

Both CD and SpA are termed complex genetic traits, because many genes are probably involved

in pathogenesis, and environmental factors have a substantial influence on the outcome of the disease. Evidence exists for a common genetic risk factor in the development of subclinical intestinal inflammation in first-degree relatives of patients with ankylosing spondylitis, which is the prototype of SpA (Bjarnason *et al.* 2003). Furthermore, we found that *CARD15*, which was the first CD susceptibility gene identified, is associated with chronic subclinical inflammation in patients with SpA (Laukens *et al.* 2004). In this regard, SpA patients can serve as a unique model for early CD.

In order to determine whether the association between the two disorders occurs not only at the clinical but also at the transcriptome level, we compared global gene expression in non-inflamed colon biopsies from SpA and CD patients. We propose that it is possible to identify a set of genes that distinguish CD patients and SpA patients with a history of chronic gut inflammation from SpA patients without chronic gut inflammation and from controls.

METHODS

Patients, tissue collection and histological classification

Colon biopsies from CD and SpA patients and healthy controls were obtained during colonoscopy. All biopsies were taken from non-inflamed sigmoid at 30 cm. Biopsy specimens were immediately placed in RNAlater (Ambion, Cambridgeshire, UK) and frozen at -80°C until sample processing. Three biopsies were obtained from each of 34 patients diagnosed with CD according to clinical, endoscopic and histological criteria, and 20 patients diagnosed with SpA according to ESSG criteria (Dougados *et al.* 1991). Sixteen patients without clinical manifestations of CD or SpA, who were undergoing colonoscopy for colon cancer screening, were included as a control population.

Histological classification of the SpA ileum and colon was performed as in our previous studies (De Vos 1989; Mielants *et al.* 1995a; Mielants *et al.* 1995b; Mielants *et al.* 1995c; De Vos *et al.* 1996). We distinguished three classes: patients with normal histology, patients with acute inflammatory lesions, and those with chronic inflammatory lesions (Cuvelier *et al.* 1987). In acute lesions, normal architecture was well preserved. There was infiltration by neutrophils and eosinophils without a considerable increase in lymphocytes. Small superficial ulcers covered with fibrin and neutrophils overlying hyperplastic lymphoid follicles were occasionally observed. The lamina propria was edematous and hemorrhagic, and contained mainly polymorphonuclear cells. The pattern of inflammation was similar to that seen in acute self-limiting bacterial enterocolitis. The principal features of chronic lesions were crypt distortion,

atrophy of the villous surface of the mucosa, villous blunting and fusion, increased mixed cellularity, and basal lymphoid aggregates in the lamina propria. Though several biopsies were obtained from each patient, a diagnosis of chronic inflammation was made even if only one biopsy showed chronic lesions, regardless of acute or active inflammation in the other biopsies.
SpA patients who had chronic inflammation in colon and/or ileum in previous examinations were termed SpA patients with chronic gut inflammation.

RNA extraction

Total RNA was extracted from biopsies using the Qiagen Rneasy Mini Kit (Westburg BV, Leusden, The Netherlands) with on-column DNAse treatment (Qiagen). Needle homogenization was performed. RNA quality and concentration were checked on the Agilent 2100 Bioanalyzer (Agilent Technologies, Germany).

Macroarray hybridization and analysis

Colony filters containing 74,828 expressed sequence tag (EST) clones (Human UniGene collection 2, RZPD, Germany) were used as initial screen. Radioactively labeled probes were produced by incorporation of [α^{33}P]dCTP during reverse transcription of 50 μg total RNA (MMLV, Promega, Leiden, The Netherlands), using oligodT as primer. ^{33}P-cDNA probes were purified on G-50 spin columns (Amersham Biosciences, Roosendaal, The Netherlands). Hybridization was performed at 10^6 cpm/ml at 65°C for 20 hours. Images were acquired after 6, 18 and 24 hours of exposure, using a Phosphorimager system (Amersham Biosciences). Spot definition and intensity measurement was done using Visualgrid (GPC Biotech AG, Munich, Germany). The raw expression data were processed with an in-house algorithm based on MS Access. Spot intensities were corrected for the local background, followed by a quality control of spots to exclude those influenced by intense signals of adjacent spots. The detection limit for expression values above background was calculated based on the variation of the local background intensity. Constitutive genes (those that show the lowest coefficient of variation over all arrays) were used for normalization. Subsequently, quantitative measures of each clone (gene) were calculated by $\log_2$ transformation of the ratio of the mean spot intensity of CD or SpA patients to the mean spot intensity of controls.

Microarray hybridization, scanning and analysis

Construction of a focus microarray chip, probe labeling, hybridization, washing and scanning were carried out at the MicroArray Facility of the Flanders Interuniversity Institute for

Biotechnology (MAF, Leuven, Belgium). Clones selected from the macroarray screen were PCR-amplified from RZPD clones using universal M13 primers. PCR fragments were purified on MultiScreen PCR plates (Millipore, Brussels, Belgium) and resuspended in 50% DMSO at an average concentration of 100 ng/µl. The PCR products were arrayed in duplicate on Type VII silane-coated slides using a Molecular Dynamics Generation III printer (Amersham Biosciences, Buckinghamshire, UK). Total RNA (5 µg) was amplified using a modified protocol of *in vitro* transcription as described (Puskas *et al.* 2002). Five µg of the amplified RNA were Cy3- or Cy5-labeled as described at http:\\www.microarrays.be\service.htm. Arrays were scanned at 532 and 635 nm using a Generation III scanner (Amersham BioSciences). Images were analyzed with ArrayVision (Imaging Research Inc., Ontario, Canada). Each hybridization was repeated in a dye swap experiment. Spot intensities were measured, corrected for local background, and those that exceeded the background by more than two standard deviations (SD) were included. For each gene, ratios of red (Cy-5) to green (Cy-3) intensities (I) were calculated and normalized via a Lowess Fit of the $\log_2$ ratios [$\log_2$(ICy-5 / ICy-3)] over the $\log_2$ total intensity [$\log_2$(ICy-5 × ICy-3)].

For comparing the microarray datasets, a mixture of RNA from 5 CD patients, 5 SpA patients and 5 controls was used as reference RNA. This reference sample provides a positive hybridization signal at each probe element on the microarray, which is essential when calculating and comparing fluorescence ratios. The data were imported into GeneMaths® XT (Applied Maths, St-Martens-Latem, Belgium). Weighted mean ratios and their corresponding error (pixel SD) were calculated from the dye swap. Data were normalized over all arrays, and missing values were imputed using k-nearest neighbor algorithm (20 neighbors). GeneMaths® XT was used to perform all subsequent supervised and unsupervised analyses.

Statistics

All *P*-values chosen for cut-off are subjective.

RESULTS

Design of the custom microarray

In order to provide a practical and cost effective tool for conducting a large number of hybridizations, a self-designed focus microarray chip was constructed specifically for studying colonic gene expression in SpA and CD. To accomplish this, a genome-wide survey of gene expression in colon biopsies of 4 CD patients, 4 SpA patients and 6 controls was conducted

using high-density nylon arrays containing 74,828 cDNA sequences (Table 1, macroarrays). Spots that showed aberrant morphology, encompassed variation in replicates or were impaired because of over shining (characteristic of radioactive signals) were filtered out and considered as clones lost through experimental error. To select for clones that are differentially expressed in CD or SPA patients, we arbitrarily selected for those that have a log_2 transformed mean ratio of less than -0.6 or more than +0.6 (1.5 fold down- or up-regulated). Genes that might be differentially expressed between groups (control versus CD or control versus SpA) were identified using a simple algorithm based on the t-test ($P<0.05$) and F-values ($P<0.05$) as selection criteria, providing that at least three consistent intensity values were present in each group. F-values were chosen for selection because we assumed that differences in variances within groups can be important. A total of 2,652 clones were identified as "potentially differentially expressed". These genes, together with 4,127 ESTs lost through experimental error and which might include, beside control ESTs, additional differentially expressed genes, were used to produce a glass-based microarray platform. This allowed us to screen more patients in a more accurate and sensitive manner.

Clustering of unfiltered data

We hybridized an independent cohort of 15 SpA, 11 CD and 10 control patients to the focus microarray (Table 1, microarrays). Unsupervised clustering (without prior knowledge of groups) using all genes revealed no clustering with respect to disease or phenotype (e.g. type of intestinal inflammation). The inability to find discriminatory genes using unfiltered data is perhaps not too surprising, as we are analyzing the steady state transcriptome in non-inflamed tissue samples of complex inflammatory diseases. Subtle differences in only a few genes are lost in the vast number of random variations. The problem of detecting differentially expressed genes can be overcome by performing supervised clustering. To this end we divided the patients into four main groups: CD, SpA with chronic gut inflammation, SpA without chronic gut inflammation, and controls. Discriminant analysis can reduce *N*-dimensional data into a more visual 2-D or 3-D plot, with prior knowledge of groups (Figure 1). With this approach, the above-defined groups became clearly separated, indicating that our full dataset contains genes that can differentiate between these disease states.

Table 1 Study population

	Diagnosis	Sample	Sex	Age	SpA	Gut Histology	Clinical CD	CD Location	Medication
macroarrays	Control	1	F	59	no	normal	no		-
		2	F	45	no	normal	no		-
		3	F	55	no	normal	no		-
		4	M	30	no	normal	no		-
		5	F	58	no	normal	no		-
		6	M	40	no	normal	no		-
	CD	7	M	21	no	chronic	yes	IC	-
		8	F	48	no	chronic	yes	IC	5-ASA
		9	F	41	yes	chronic	yes	C	-
		10	F	23	no	chronic	yes	C	AZA
	SpA	11	M	47	yes	normal	no		-
		12	F	85	yes	normal	no		-
		13	F	60	yes	normal	no		-
		14	F	43	yes	chronic	yes		-
microarrays	Control	1	M	54	no	normal	no		-
		2	F	64	no	normal	no		-
		3	F	72	no	normal	no		-
		4	M	51	no	normal	no		-
		5	F	21	no	normal	no		-
		6	F	68	no	normal	no		-
		7	F	32	no	normal	no		-
		8	F	73	no	normal	no		-
		9	F	66	no	normal	no		-
		10	M	76	no	normal	no		-
	CD	11	F	23	no	chronic	yes	IC	5-ASA
		12	F	39	yes	chronic	yes	C	-
		13	M	51	no	chronic	yes	C	-
		14	M	43	no	chronic	yes	I	-
		15	M	36	no	chronic	yes	IC	-
		16	F	27	no	chronic	yes	IC	-
		17	F	46	yes	chronic	yes	IC	AZA
		18	F	23	no	chronic	yes	IC	-
		19	M	19	no	chronic	yes	IC	-
		20	F	26	no	chronic	yes	IC	-
		21	F	36	no	chronic	yes	IC	-
	SpA	22	M	40	AS periph	acute	no		-
		23	M	29	AS periph	normal	no		NSAID
		24	M	42	AS periph	acute	no		-
		25	M	31	AS periph	acute	no		NSAID
		26	M	76	uSPA	normal	no		steroids
		27	F	28	AS ax	normal	no		NSAID
		28	M	58	AS periph	normal	no		-
		29	M	38	AS ax	normal	no		sulfa
		30	F	49	AS periph	chronic	no		sulfa
		31	M	49	AS periph	chronic	no		-
		32	M	45	AS periph	chronic	no		sulfa
		33	M	29	uSPA	acute	no		sulfa+NSAID
		34	M	48	AS periph	normal	no		NSAID
		35	F	44	AS periph	chronic	yes	IC	sulfa+AZA
		36	M	36	AS ax	normal	no		sulfa+NSAID

UspA = undifferentiated SpA; AS ax = ankylosing spondylitis with only axial involvement; AS periph = ankylosing spondylitis with peripheral involvement; histology of SpA patients is a historical classification; IC = ileocolonic; I = ileal involvement only; C = colonic involvement only; 5-ASA = 5-aminosalicylates; AZA = azathioprine; NSAID = non-steroidal anti-inflammatory drugs; sulfa = sulfasalazine

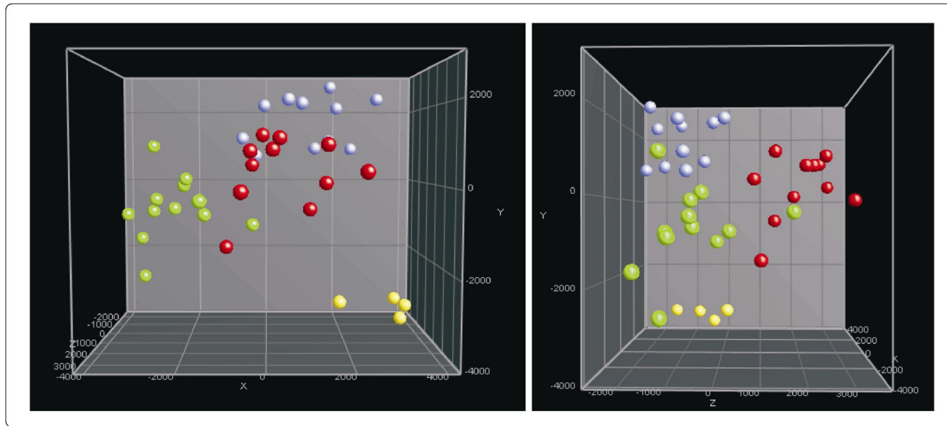


Figure 1 Discriminant analysis of all patients using unfiltered data, illustrated in two directions. Four groups are clearly separated. CD (green), SpA without chronic inflammation (blue), SpA with chronic inflammation (yellow) and controls (red).

Identification of genes whose differential regulation is common to both SpA with chronic gut inflammation and CD

By using an independent t-test, we identified 123 genes that are expressed differentially between CD and control (P<0.01). With this set we were unable to discriminate SpA patients from controls, although three out of four of the SpA patients with chronic gut inflammation clustered together, indicating the presence of alterations similar to those observed in CD. Thus it was logical to screen for genes modulated commonly in CD and control on one hand, and SpA, SpA with chronic gut inflammation and control on the other hand. In order to include a larger number of genes in this analysis, the statistical significance level was lowered from P<0.01 to P<0.05. This led to the identification of two sets of genes whose expression pattern discriminates CD from control (P<0.05, N=630) and SpA from control (P<0.05, N=464). The latter significance level was determined by ANOVA, in which SpA patients with chronic gut inflammation were defined as a distinct group. The set of 95 genes that are differentially expressed in both CD and SpA could distinguish the three disease groups (Figure 2A, Table 2). In addition, SpA patients with chronic gut inflammation cluster together and are more related to the CD cluster than to the control/SpA cluster, but they remain a separate entity (Figure 2A). Principal component analysis using this set of 95 genes, another way of representing the data, clearly discriminates our patient groups (Figure 2B). Based on the identification of a set of genes involved in CD and also implicated in SpA, we attempted to identify genes that might render these individuals more susceptible to develop CD.

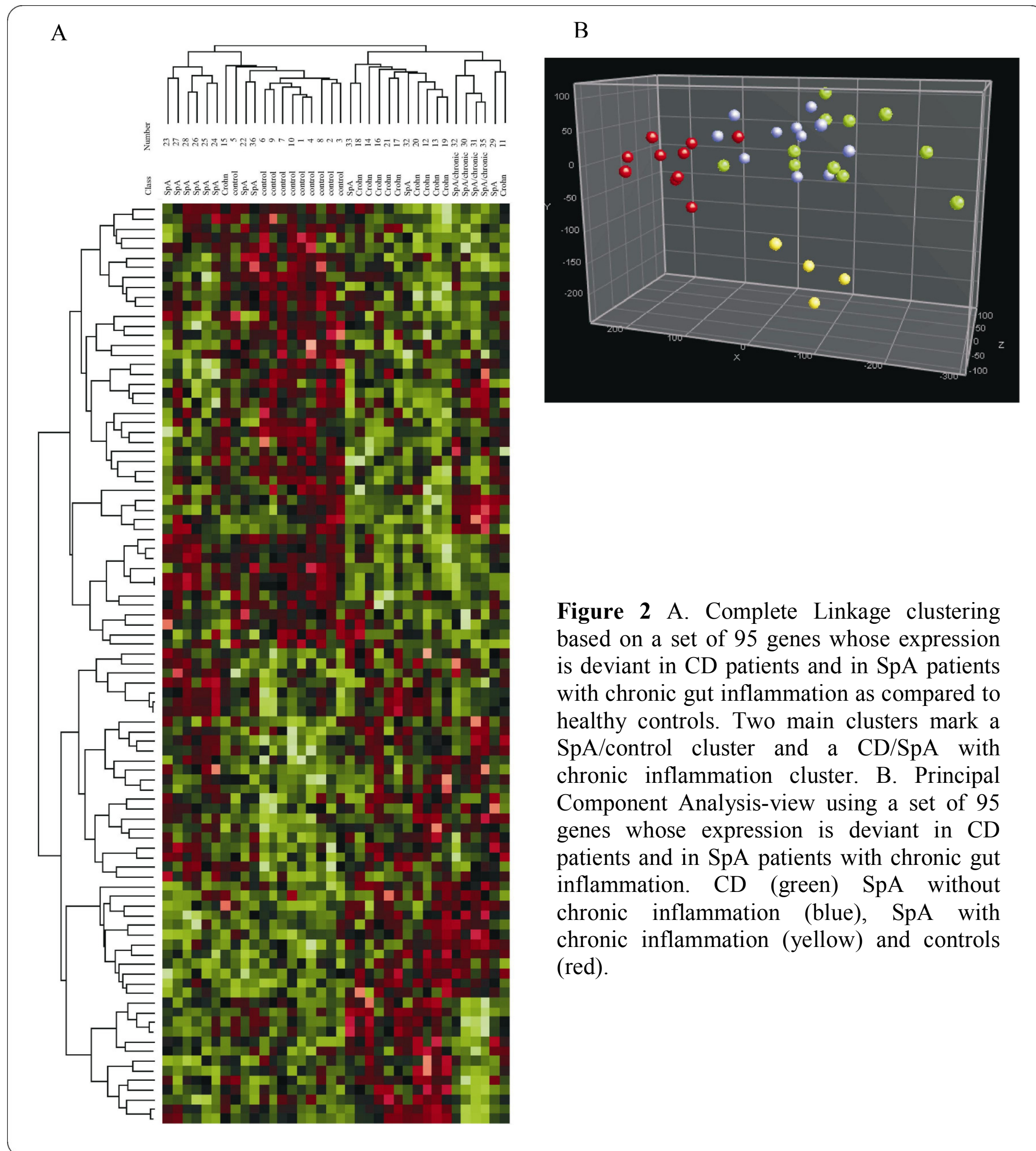


Figure 2 A. Complete Linkage clustering based on a set of 95 genes whose expression is deviant in CD patients and in SpA patients with chronic gut inflammation as compared to healthy controls. Two main clusters mark a SpA/control cluster and a CD/SpA with chronic inflammation cluster. B. Principal Component Analysis-view using a set of 95 genes whose expression is deviant in CD patients and in SpA patients with chronic gut inflammation. CD (green) SpA without chronic inflammation (blue), SpA with chronic inflammation (yellow) and controls (red).

Genes within the CD/ SpA chronic cluster

Genes whose expression is aberrant in both CD and SpA with chronic gut inflammation are represented in Table 2 (↑: up-regulated; ↓: down-regulated, $P<0.05$). Among them, two genes had already been described in the context of CD. Acyl-coenzyme A oxidase 1 (ACOX1), which is the first enzyme of the fatty acid beta-oxidation pathway, donates electrons directly to molecular oxygen, thereby producing hydrogen peroxide. The enzymatic activity of ACOX1 was diminished in both inflamed and non-inflamed areas in CD (Aimone-Gastin *et al.* 1994). Our observation of ACOX1 transcript down-regulation corroborates this report, and indicates a fault at the level of transcription or mRNA stability.

Table 2 Ninety-five EST's that cluster CD and SpA patients with chronic gut inflammation

Accession	Gene description	Chromosomal location	Expression in CD/SpA chronic	Reference
H81904	acetyl-Coenzyme A synthetase 2 (ADP forming)	20q11.22	↓	
AI758270			↓	
R20596	chromosome 6 open reading frame 32	6p22.3-p21.32		(Barmada *et al.* 2004)
AI349525	hypothetical protein FLJ34790	17p13.1	↓	
AI445844	pyruvate dehydrogenase kinase, 3	Xq22	↓	(van Heel *et al.* 2004)
H56656		5q12.3	↓	
N22829	GDP dissociation inhibitor 2	10p15	↓	
AI274555	TatD DNase domain containing	8q24.13		(Vermeire *et al.* 2004)
H61436	sorting nexin 24	5q23.2		
H15751	poly (ADP-ribose) polymerase family, member 16	15q22.31	↓	
H87107	ring finger protein 5	6p21.3	↓	(Rioux *et al.* 2000)
AI680066				
AI191504	adenylate kinase 7	14q32.2	↓	
R17390	sorting nexin 15	11q12		
AA812701	hypothetical protein LOC340178	6q27		
AA825971		2q31		
AA291593	solute carrier family 39 (metal ion transporter) member 11	17q24.3-q25.1		
	myoglobin	22q13.1		(Barmada *et al.* 2004)
AA420968	chromosome 6 open reading frame 112	6q21		
R08643	interferon regulatory factor 2 binding protein 2	1q42.3		
T83584	mannosidase, beta A, lysosomal	4q22-q25		(Vermeire *et al.* 2004)
AA481482	E74-like factor 3 (ets domain transcription factor, epithelial-specific)	1q32.2		
T78477	zinc finger protein 297	6p21.3		(Rioux *et al.* 2000)
H27674	DAB2 interacting protein	9q33.1-q33.3		
R13194	BRF1 homolog,	14q		
R98758	hypothetical protein FLJ32731	8p11.1		
H96900	MGC2747	19p13.11		
N62199	kinesin family member 1B	1p36.2	↓	(Cho *et al.* 2000)
R51052	S-phase response (cyclin-related)	1q42.11-q42.3	↓	
AI286348	SH3 domain containing ring finger 2	5q32		
R17478	proliferation-inducing protein 13	1q25		
AI916359	secretory pathway calcium ATPase 2	16q24.1		
H19034	bladder cancer associated protein	20q11.2-q12		
N39050	cDNA FLJ39784 fis, clone SPLEN2002314			
R13544	F-box and leucine-rich repeat protein 3A	13q22	↓	
AI286163			↓	
R07321	Acyl-Coenzyme A oxidase 1, palmitoyl	17q24-17q25	↓	
R12168	staufen, RNA binding protein, homolog 2 (Drosophila)	8q13-q21.1	↓	
H85472	N-myc downstream regulated gene 1	8q24.3	↓	
AA513663			↓	
R97820	zinc finger CCCH type, antiviral 1	7q34	↓	
N62837	leukocyte immunoglobulin-like receptor, subfamily A 4	19q13.4	↓	
AA019615	ubiquitin carboxyl-terminal hydrolase L5	1q32	↓	
AI809092			↓	
AA758064				
AI219353	cancer susceptibility candidate 2	10q26.11	↓	
AI698801				
R12632	hematological and neurological expressed 1	17q25.1		
AA468418				
	histone 1, H2ac	6p21.3		(Barmada *et al.* 2004)
H23734	thymosin, beta 4, X-linked	Xq21.3-q22		
H75688	sorting nexin 7	1p21.3		
N63669	triple functional domain (PTPRF interacting)	5p15.1-p14		
AI819016			↑	
AI831827	ADP-ribosylation factor guanine nucleotide-exchange factor 2	20q13.13		
AI923299	hypothetical LOC158730	Xp21.1	↑	
AI286257	zinc finger protein 263	16p13.3		
AI863417	runt-related transcription factor 1	21q22.3	↑	

***Table 2** continued*

Accession	Gene description	Chromosomal location	Expression in CD/SpA chronic	Reference
AI274438			↑	
AI289775	weakly similar to 0412263A Ig G2 pFc' PIG Gm		↑	
AI345173				
AI928384				
AA490519	clone IMAGE:5286005			
AI696951	FLJ21763 fis			
AI339536	IMAGE:5286812, mRNA			
AI245190				
AI261428				
AI809310	hect domain and RCC1 (CHC1)-like domain 1	15q22		
AI831536	interleukin 12 receptor, beta 2	1p31.3-p31.2		
T79944			↑	
AA480677	IMAGE:4865533		↑	
AA009461	MUF1 protein	1p34.1	↑	(Vermeire *et al.* 2004)
R61783	synaptogyrin 2	17q25.3	↑	
AI282992	glutathione peroxidase 2 (gastrointestinal)	14q24.1	↑	
W65310				
AA588676	PDZ and LIM domain 3	10q22.3-q23.2		
AI811147			↑	
AI357718				
N58523				
W74496			↑	
AI352320			↑	
AI280708	thymidylate kinase family LPS-inducible member	2p25.2		
AI815599	proteasome (prosome, macropain) 26S subunit, non-ATPase, 11	17q11.2		
AI718161				
N98921	tropomyosin 3	1q21.2		(Vermeire *et al.* 2004)
AI367201	HSPC135 protein	3q13.2		
AA425545	RP42 homolog	3q26.3		
H44213	hypothetical gene supported by BC031266			
N58936	nuclear receptor coactivator 7	6q22.32		
N90208	hypothetical protein FLJ32440	8q24.13		(Vermeire *et al.* 2004)
AA427403	interferon induced transmembrane protein 1 (9-27)	11p15.5		
R25648	phosphatidylinositol 4-kinase type-II beta	4p15.2		
N93265	superoxide dismutase 1, soluble (amyotrophic lateral sclerosis)	21q22.11		
H99865	nuclear receptor subfamily 2, group F, member 2	15q26		

Genes with aberrant expression in both CD and SpA patients with chronic gut inflammation (↓ down-regulated and ↑ up-regulated as compared to the control and SpA population ($P<0.05$). Genetic markers are cited for genes that are located within or near (5 cM) one of the CD loci

Glutathione peroxidase 2 (gastrointestinal glutathione peroxidase, giGSH-Px) is one of the four types of selenium-dependent glutathione peroxidases. Its exclusive expression in the gastrointestinal tract indicates that it functions as a barrier against the absorption of dietary hydroperoxides, and as protection against damage from endogenously formed hydroxyl peroxides. Its activity is increased in ulcerative colitis patients in the active and in the remission stages (Beno *et al.* 1997). In CD, plasma levels of giGSH-Px are increased (Tuzun *et al.* 2002). We found that this gene is over-expressed in normal colon tissue in CD and SpA patients with a history of chronic gut inflammation, and so it can act as a marker expressed at non-pathological sites in the intestine in CD and CD susceptible SpA patients.

CONCLUSION

Clinical study of intestinal abnormalities in SpA patients has previously relied on cytokine profiles and immunological alterations. In addition to analyzing individual proteins, genome-wide transcript profiles can be analyzed by microarrays. Global gene expression analysis in non-inflamed colon tissue was used to find genes that are differentially expressed in both CD patients and SpA patients with a history of chronic gut inflammation. Previous studies of gene expression in IBD have focused on biopsies of actively inflamed tissues (Lawrance *et al.* 2001; Dooley *et al.* 2004; Stoll *et al.* 2004). The use of samples from non-inflamed areas from CD patients offers the possibility of identifying early markers for CD, which would permit prediction of the evolution to CD in SpA patients. Moreover, changes in the expression of genes that are regulated during inflammation would be more prominent than the subtle alterations in non-inflammatory genes, although it cannot be ruled out that this procedure will also pick up genes whose differential expression is a consequence and not a cause of the disease. Additionally, looking at basal gene expression may allow us to take into account genetic influences, since gene expression is highly heritable (Morley *et al.* 2004). Therefore, future studies on markers for CD should concentrate primarily on genes that are located near one of the known loci for CD (Table 2). Genes located within a region linked to CD or IBD in general (if multipoint linkage was performed), or within five centimorgan of the markers that are linked to CD or IBD (in case of two point linkage) should be considered first. Using a model for early CD when identifying CD susceptibility genes can circumvent the heterogeneity of the disease, because probably only a very small number of CD genes will be implicated in SpA.

Array analysis is a rapid procedure for studying the expression of many genes in no more than several samples. Because the number of samples is limited, and the number of genes explored is usually large, false-positive results will obviously arise. Nevertheless, array analysis enables us to explore gene expression with different computational tools. To confirm the importance of a set of genes associated with a phenotype, complementary techniques such as quantitative PCR (qPCR) are mandatory. Thus, arrays are not simply a way to find single differentially regulated genes; they can be used to compare global gene expression in distinct groups.

We show that SpA patients have an aberrant gene expression profile in comparison to healthy controls, indicating that alteration of gene expression in the colon of SpA patients is a biologically relevant concept. We identified a set of genes that are differentially expressed in both CD and SpA patients who are at higher risk of developing CD. Based on the expression of these 95 genes, SpA patients with subclinical chronic gut inflammation cluster with CD,

confirming the clinical association between the two inflammatory disorders. We also suggest a number of candidate genes for mutation screening. We are currently verifying a selection of genes by qPCR, and exploring the involvement of genes that are differentially expressed in both CD and SpA patients with a history of chronic gut inflammation, in order to find early (genetic) markers for CD in SpA patients.

REFERENCES

Aimone-Gastin, I., S. Cable, J. M. Keller*, et al.* (1994). "Studies on peroxisomes of colonic mucosa in Crohn's disease." Dig Dis Sci **39**(10): 2177-85.

Baeten, D., P. Demetter, C. A. Cuvelier*, et al.* (2002). "Macrophages expressing the scavenger receptor CD163: a link between immune alterations of the gut and synovial inflammation in spondyloarthropathy." J Pathol **196**(3): 343-50.

Barmada, M. M., S. R. Brant, D. L. Nicolae*, et al.* (2004). "A genome scan in 260 inflammatory bowel disease-affected relative pairs." Inflamm Bowel Dis. **10**(5): 513-20.

Beno, I., M. Staruchova and K. Volkovova (1997). "Ulcerative colitis: activity of antioxidant enzymes of the colonic mucosa." Presse Med **26**(31): 1474-7.

Bjarnason, I., K. O. Helgason, A. J. Geirsson*, et al.* (2003). "Subclinical intestinal inflammation and sacroiliac changes in relatives of patients with ankylosing spondylitis." Gastroenterology **125**(6): 1598-605.

Cho, J. H., D. L. Nicolae, R. Ramos*, et al.* (2000). "Linkage and linkage disequilibrium in chromosome band 1p36 in American Chaldeans with inflammatory bowel disease." Hum Mol Genet **9**(9): 1425-32.

Cuvelier, C., C. Barbatis, H. Mielants*, et al.* (1987). "Histopathology of intestinal inflammation related to reactive arthritis." Gut **28**(4): 394-401.

Davis, P., A. B. Thomson and B. C. Lentle (1978). "Quantitative sacroiliac scintigraphy in patients with Crohn's disease." Arthritis Rheum **21**(2): 234-7.

De Vos, M. (1989). "Ileocolonoscopy in seronegative spondylarthropathy." Gastroenterology **96**: 339-344.

De Vos, M., H. Mielants, C. Cuvelier*, et al.* (1996). "Long-term evolution of gut inflammation in patients with spondyloarthropathy." Gastroenterology **110**(6): 1696-703.

Demetter, P., D. Baeten, F. De Keyser*, et al.* (2000). "Subclinical gut inflammation in spondyloarthropathy patients is associated with upregulation of the E-cadherin/catenin complex." Ann Rheum Dis **59**(3): 211-6.

Demetter, P., J. A. Van Huysse, F. De Keyser*, et al.* (2002). "Increase in lymphoid follicles and leukocyte adhesion molecules emphasizes a role for the gut in spondyloarthropathy

pathogenesis." J Pathol **198**(4): 517-22.

Dooley, T. P., E. V. Curto, S. P. Reddy*, et al.* (2004). "Regulation of gene expression in inflammatory bowel disease and correlation with IBD drugs: screening by DNA microarrays." Inflamm Bowel Dis **10**(1): 1-14.

Dougados, M., S. van der Linden, R. Juhlin*, et al.* (1991). "The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy." Arthritis Rheum **34**(10): 1218-27.

Elewaut, D., F. De Keyser, C. Cuvelier*, et al.* (1998a). "Distinctive activated cellular subsets in colon from patients with Crohn's disease and ulcerative colitis." Scand J Gastroenterol **33**(7): 743-8.

Elewaut, D., F. De Keyser, F. Van Den Bosch*, et al.* (1998b). "Enrichment of T cells carrying beta7 integrins in inflamed synovial tissue from patients with early spondyloarthropathy, compared to rheumatoid arthritis." J Rheumatol **25**(10): 1932-7.

Laukens, D., H. Peeters, D. Marichal*, et al.* (2004). "*CARD15* gene polymorphisms in patients with spondyloarthropathies identify a specific phenotype previously related to Crohn's disease." Ann Rheum Dis. **64**(6): 930-935.

Lawrance, I. C., C. Fiocchi and S. Chakravarti (2001). "Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes." Hum Mol Genet **10**(5): 445-56.

Mielants, H., E. M. Veys, C. Cuvelier*, et al.* (1995a). "The evolution of spondyloarthropathies in relation to gut histology. II. Histological aspects." J Rheumatol **22**(12): 2273-8.

Mielants, H., E. M. Veys, C. Cuvelier*, et al.* (1995b). "The evolution of spondyloarthropathies in relation to gut histology. III. Relation between gut and joint." J Rheumatol **22**(12): 2279-84.

Mielants, H., E. M. Veys, M. De Vos*, et al.* (1995c). "The evolution of spondyloarthropathies in relation to gut histology. I. Clinical aspects." J Rheumatol **22**(12): 2266-72.

Morley, M., C. M. Molony, T. M. Weber*, et al.* (2004). "Genetic analysis of genome-wide variation in human gene expression." Nature **430**(7001): 743-7.

Puskas, L. G., A. Zvara, L. Hackler, Jr.*, et al.* (2002). "RNA amplification results in reproducible microarray data with slight ratio bias." Biotechniques **32**(6): 1330-4, 1336, 1338, 1340.

Rioux, J. D., M. S. Silverberg, M. J. Daly*, et al.* (2000). "Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci." Am J Hum Genet. **66**(6): 1863-70.

Scott, W. W., Jr., E. K. Fishman, J. E. Kuhlman*, et al.* (1990). "Computed tomography evaluation of the sacroiliac joints in Crohn disease. Radiologic/clinical correlation." Skeletal Radiol **19**(3): 207-10.

Stoll, M., B. Corneliussen, C. M. Costello*, et al.* (2004). "Genetic variation in DLG5 is

associated with inflammatory bowel disease." Nat Genet. **36**(5): 476-80.

Tuzun, A., A. Erdil, V. Inal*, et al.* (2002). "Oxidative stress and antioxidant capacity in patients with inflammatory bowel disease." Clin Biochem **35**(7): 569-72.

van Heel, D. A., S. A. Fisher, A. Kirby*, et al.* (2004). "Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs." Hum Mol Genet **13**(7): 763-70.

Vermeire, S., P. Rutgeerts, K. Van Steen*, et al.* (2004). "Genome wide scan in a Flemish inflammatory bowel disease population: support for the IBD4 locus, population heterogeneity, and epistasis." Gut **53**(7): 980-6.

Future prospects

THE IMPORTANCE OF GENE EXPRESSION PROFILING IN MAPPING COMPLEX GENETIC TRAITS

Debby Laukens

Department of Molecular Biomedical Research, Ghent University and Flanders Interuniversity Institute for Biotechnology (VIB), Gent, Belgium

INTRODUCTION

Microarrays are the ideal tool to study many thousands of genes at once. However, since their widespread use, little has been accomplished, except for simply identifying differentially expressed genes. Usually, the processing and interpretation of the huge amount of data is difficult. Another application based on microarrays is the extraction of new classification algorithms. However, this demands a lot of samples, and is thus not cost-effective. Additionally, microarrays are used for the identification of pathways. However, many genes will be missed, because not all expression changes will occur at one specific time-point. Together with the statistical set backs, these are the reasons why studies on microarrays are increasingly criticized. They seem to be merely applicable as an initial screen for differentially expressed genes, and absolutely require additional verification. Moreover, it was recently postulated that the use of microarrays to study transcription profiling in the future might become less important because "it focuses on an intermediate" (Hoheisel 2006). Indeed, RNA might be too far removed from the actual cellular effectors to be used for diagnostic purposes, unless regulatory RNAs, such as microRNAs are targeted (He *et al.* 2004). We believe, however, that the RNA expression of genes is crucial in the development of diseases, and is probably a major determinant for common phenotypic characteristics, such as height or intelligence. Therefore, transmitted gene expression patterns associated with complex genetic diseases should provide useful information, but up to now, this has been underestimated in genetic studies. Here, we consider the use of microarrays in the identification of susceptibility genes, and we propose a novel method based on the integration of microarray analysis in family pedigrees.

NATURAL VARIATION AND SUSCEPTIBILITY

Natural genetic variation at the DNA level greatly influences the basal expression of genes or their targets. Analysis of microarrays to investigate the genetic circuitry that regulates expression has first been used in yeast (DeRisi *et al.* 1997). It is a favourable organism to study gene expression, since the genes are easy to recognize, and the *cis* regulatory elements are compact and close to the transcription units. Certain yeast mutants have been characterized by their expression pattern and sets of genes could be grouped and rationalized by identification of upstream regulatory sequences. In mice, differential gene expression induced by single-gene effects in complex tissues has been studied (Callow *et al.* 2000; Aronow *et al.* 2001). In contrast to the exploration of functional consequences of a defined genetic difference, these approaches have also been used in combination with traditional quantitative trait locus (QTL) mapping techniques. Indeed, Karp and co-workers have successfully identified complement factor 5 (C5) as a susceptibility gene in a murine model for allergic asthma (Karp *et al.* 2000). C5 appeared to be the only gene that was differentially expressed between susceptible and non-susceptible backcross mice as well as located near one of the previously identified QTL intervals. Similarly, Aitman and colleagues have combined a congenic strain strategy with micorarray expression analysis in the spontaneously hypertensive rat, a model for human insulin-resistance type 2 diabetes (Aitman *et al.* 1999). Hence, CD36 was pursued as a candidate for the observed QTL at chromosome 4, and sequencing revealed multiple variants caused by unequal genomic recombination of a duplicated ancestral section.

In humans, little has been studied on gene expression and susceptibility. However, remarkable support comes from the study of allele specific variation in gene expression (Lo *et al.* 2003). Among 602 heterozygous genes that contained transcribed single nucleotide polymorphisms, 54% displayed at least a twofold preferential expression of one allele. This allelic variation was shown to be transmitted by mendelian inheritance (Yan *et al.* 2002). Furthermore, the variance in expression level is highest among unrelated individuals, and smallest between members of monozygotic twin pairs, suggesting that germ-line differences contribute to variation in gene expression (Cheung *et al.* 2003; Correa *et al.* 2004). Four studies illustrated that basal gene expression levels are highly heritable (Watts *et al.* 2002; Cheung *et al.* 2003; Correa *et al.* 2004; Pastinen *et al.* 2004). Watts and co-workers have defined a unique expression phenotype to carriers of a recessive disease called ataxia telangiectasia. Although many studies have addressed the inheritance of single-gene mutations, so far the genome-wide inheritable expression in human disease has not been studied (Cheung *et al.* 2002). Moreover, Morley and colleagues

have shown that these quantitative traits allow the genetic mapping of determinants that contribute to variation in human gene expression (Morley *et al.* 2004). Taken together, these results demonstrate the feasibility and utility of exploring genome-wide gene expression in the study of susceptibility to human diseases.

Polymorphisms and gene expression

How can polymorphisms affect gene expression? First of all, mutations in coding sequences often result in an altered half-life of the RNA. Secondly, polymorphisms in untranslated regions, or introns can have modest or substantial influences on gene expression. Not only transcription factor binding sites or enhancer sites can be changed, but also the stability of the RNA is depending on seemingly "junk" sequences. Next to changes in *cis* acting elements, polymorphisms in *trans* acting factors could influence the expression of downstream genes.

The small variations in DNA sequences between individuals have been suggested as being responsible for inter-individual phenotypic variation. That is why common traits or phenotypic characteristics are supposed to result from natural variation. In fact, all forms of life, including life-threatening diseases and normal phenotypes, probably share the same basic idea: it is merely the impact of the polymorphism that determines the functional consequence (Figure 1). If the impact is low, such as small phenotypic changes, the polymorphism will not be eradicated by natural selection. On the other hand, if the polymorphism is causing a severe dysfunction of a protein, it will be selected out. That is probably one of the reasons why some relatively harmless traits such as allergies are common.

Epigenetics and gene expression

Unfortunately, the classical genetic ideas are complicated by new and slowly accepted mechanisms. The concept of the all-determining DNA sequence is not complete. It appears that the environment not only affects your own life, but also the life of your progeny. It is becoming widely accepted that genes have some sort of "memory". The way this memory works is explained by epigenetic mechanisms (literally: on top of genetics), acting by switching on and of the expression of genes. This also gives a new dimension as to what "environmental factors" are in genetics. Are they the factors that the person goes through at birth, during adolescence or perhaps that his ancestors were confronted with? Epigenetic effects can act on larger regions, influencing a number of genes within the same chromosomal region. Moreover, these effects are not fixed over time within the same individual. Today, two basic epigenetic mechanisms are known: DNA methylation and histone modifications. These modifications act on the expression

of genes by interfering with chromatin packaging or the physical binding of *trans* factors to the DNA strand. Furthermore, at least for methylation, evidence exists that it is genetically transmitted.

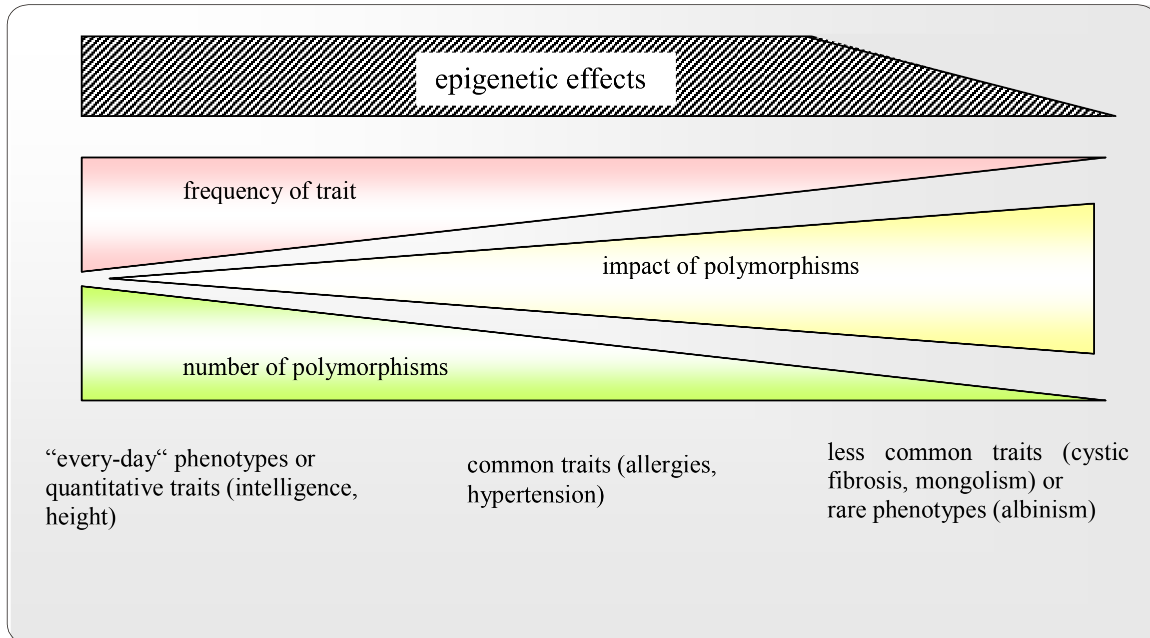


Figure 1 Natural genetic variation determines every-day life: from life-threatening diseases to the way we look.

FAMILY-BASED EXPRESSION PROFILING IN CROHN'S DISEASE

As stated above, susceptibility to CD most likely results from several variations in DNA sequence. Some of them will alter the gene product, but they may also have profound as well as subtle consequences on gene expression. On top, inherited epigenetic mechanisms will contribute to differential expression. Instead of analyzing gene expression in randomly selected CD patients, we suggest using multiple affected families to screen for sets of genes that are differentially expressed, e.g. in blood, between affected and non-affected family members (Figure 2). This way, the problem of genetic heterogeneity between random CD patients is dramatically reduced. Specific sets of genes can be identified that are transmitted only to patients with CD, and are thus ideal candidates to associate to disease. In combination with linkage data of the families, we would gain substantial evidence for an actual expression signature for CD, in which genes will have to be characterized. In the next section we briefly describe the experimental set-up of such a study.

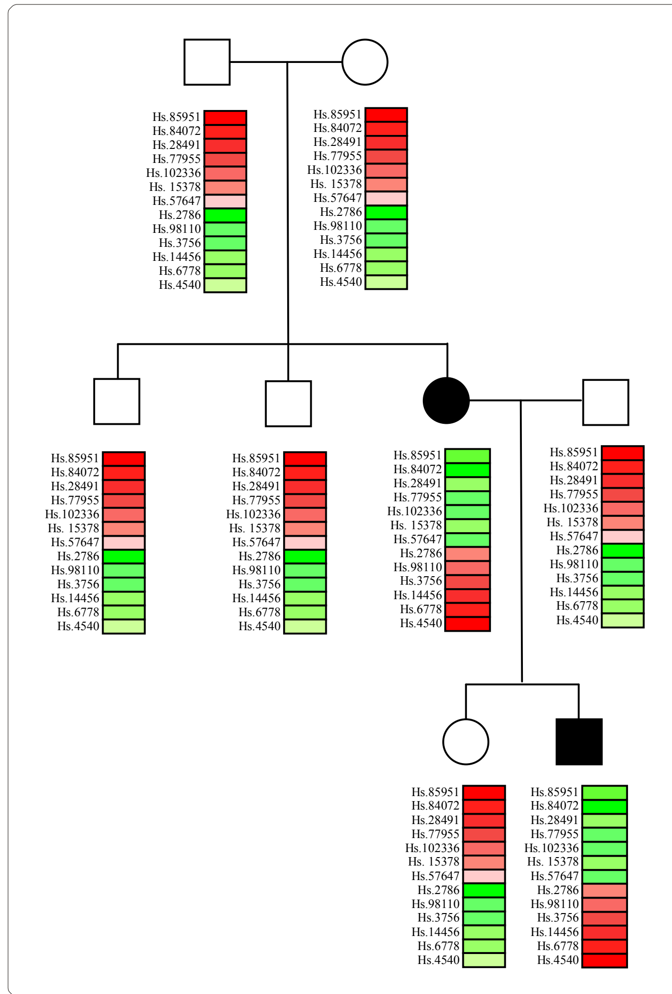


Figure 2 Simplified and theoretic example of a family pedigree used to study the transmission of expression signatures.

Design and methodologies

Gathering of appropriate families/patients and blood sampling. Patients and relatives will be asked to participate in the study. Families will be evaluated based on the magnitude, number of affected individuals, and genetic information. Patients within these families will be selected only after diagnosis is well defined, and non-affected patients should have reached a certain age, to reduce false negatives.

Collected fresh whole blood will be immediately treated with the RiboPure-Blood Kit (Ambion). This way, samples are safe to transport at ambient temperature. The majority of RNA will originate from white blood cells (WBC). The RiboPure-Blood Kit was designed to minimize contamination with heme, protein and genomic DNA during RNA extraction of whole blood. Therefore, there is no need to select for WBC before extraction, reducing experimental heterogeneity. It has been reported that some features of variation in expression in blood samples are associated with cellular composition, gender, age and time of the day, so care must be taken

in this regard (Whitney *et al.* 2003). To rule out such random effects, sampling will be done at two time points per individual, and expression profiling will be carried out in duplicate. In addition, variations observed among patients are likely to be significantly greater than the background variation in normal gene expression.

RNA extraction and microarray screening. Total RNA will be isolated within 1 day, and stored until it can be processed for RNA amplification and hybridization. We choose to perform a genome-wide screen and use the Affymetrix platform (Human Genome U133 Plus 2.0 Set, containing 47,000 transcripts including 38,500 well-characterized genes). These chips contain oligonucleotide probes, designed to maximize sensitivity, specificity, and reproducibility, allowing consistent discrimination between specific and background signals, and between closely related target sequences.

Data analysis. The raw expression data will be imported in a highly developed software program called GeneMaths® XT (Applied Maths, BVBA).
In first instance we will concentrate on those genes that are annotated and modulated between affected and non-affected individuals, base on a simple P-value cut-off. Differentially regulated genes will belong to one of the following main categories:

- false positives (due to limited statistical power)
- secondary to disease (this should be minimized, since we are not looking at the site of pathology, and genes should be at their steady-state level. In addition samples are analyzed at two time-points per individual)
- genetic variation in *trans*-acting genes
- genetic variation in *cis*-acting elements

An example of the concept and analysis of *cis* and *trans* regulators of expression in human genetics has been anticipated (Morley *et al.* 2004). Genes that are located within regions with most significant evidence of linkage are potentially those that are regulated in *cis*. Based on the information of the genes we will have to decide on which genes to focus.
Genes that are not located within these regions can still be regulated in *trans*. Extra evidence will have to be gathered to decide if and which genes might be interesting to pursue. For example, unsupervised clustering can reveal genes that are commonly regulated. Promoter analysis will be

conducted to search for common binding factors. Evidently, transcription factors that are located within a locus are considered as candidate genes.

In a more elaborate search for regulators, the expression phenotype can be analyzed genetically as a quantitative trait in order to identify the determinants of the variation in gene expression. The differentially expressed genes with the most evidence of linkage will be used to carry out genome-wide linkage analysis in CEPH families (Centre d'etude du polymorphisme humain) (Morley *et al.* 2004). These are large three-generation families, many of which have been genotyped. Immortalized lymphocytes are available from all the members of the pedigree. Consequently, loci can be attributed to specific expression phenotypes, and be evaluated by comparing to known CD loci.

Genetic screening. Genes that meet the criteria as discussed above, will be screened for mutations/polymorphisms in matching patients. For this, sequence analysis of exonic sequences and ~1000 nucleotides upstream of the start codon will be performed. Polymorphisms are then evaluated in an independent CD and control population to assess the population risk.

REFERENCES

Aitman, T. J., A. M. Glazier, C. A. Wallace, *et al.* (1999). "Identification of Cd36 (Fat) as an insulin-resistance gene causing defective fatty acid and glucose metabolism in hypertensive rats." Nat Genet **21**(1): 76-83.

Aronow, B. J., T. Toyokawa, A. Canning, *et al.* (2001). "Divergent transcriptional responses to independent genetic causes of cardiac hypertrophy." Physiol Genomics **6**(1): 19-28.

Callow, M. J., S. Dudoit, E. L. Gong, *et al.* (2000). "Microarray expression profiling identifies genes with altered expression in HDL-deficient mice." Genome Res **10**(12): 2022-9.

Cheung, V. G. and R. S. Spielman (2002). "The genetics of variation in gene expression." Nat Genet **32**(Suppl): 522-5.

Cheung, V. G., L. K. Conlin, T. M. Weber, *et al.* (2003). "Natural variation in human gene expression assessed in lymphoblastoid cells." Nat Genet **33**(3): 422-5.

Correa, C. R. and V. G. Cheung (2004). "Genetic variation in radiation-induced expression phenotypes." Am J Hum Genet **75**(5): 885-90.

DeRisi, J. L., V. R. Iyer and P. O. Brown (1997). "Exploring the metabolic and genetic control of gene expression on a genomic scale." Science **278**(5338): 680-6.

He, L. and G. J. Hannon (2004). "MicroRNAs: small RNAs with a big role in gene regulation." Nat Rev Genet. **5**(7): 522-31.

Hoheisel, J. D. (2006). "Microarray technology: beyond transcript profiling and genotype analysis." Nat Rev Genet. **7**(3): 200-10.

Karp, C. L., A. Grupe, E. Schadt*, et al.* (2000). "Identification of complement factor 5 as a susceptibility locus for experimental allergic asthma." Nat Immunol **1**(3): 221-6.

Lo, H. S., Z. Wang, Y. Hu*, et al.* (2003). "Allelic variation in gene expression is common in the human genome." Genome Res **13**(8): 1855-62.

Morley, M., C. M. Molony, T. M. Weber*, et al.* (2004). "Genetic analysis of genome-wide variation in human gene expression." Nature **430**(7001): 743-7.

Pastinen, T., R. Sladek, S. Gurd*, et al.* (2004). "A survey of genetic and epigenetic variation affecting human gene expression." Physiol Genomics **16**(2): 184-93.

Watts, J. A., M. Morley, J. T. Burdick*, et al.* (2002). "Gene expression phenotype in heterozygous carriers of ataxia telangiectasia." Am J Hum Genet **71**(4): 791-800.

Whitney, A. R., M. Diehn, S. J. Popper*, et al.* (2003). "Individuality and variation in gene expression patterns in human blood." Proc Natl Acad Sci U S A **100**(4): 1896-901.

Yan, H., W. Yuan, V. E. Velculescu*, et al.* (2002). "Allelic variation in human gene expression." Science **297**(5584): 1143.

Summary and discussion

The objective of this project was to use gene expression analysis in the search of genetic susceptibility genes for Crohn's disease (CD). Crohn's disease is a debilitating disease with substantial impact on the patients' quality of life. For some reason, the intestine becomes chronically inflamed, extending deep into the layers of the intestinal wall. The patient experiences periods of clinical relapses and remission. In general, the complete gastrointestinal tract can be affected, however, the inflammation is usually confined to the small bowel (ileum) or the large bowel (colon). In fact, the location of disease is one of the most accepted phenotypic classifications for CD, because it stays particularly stable throughout disease duration. Moreover, concordance rates for disease location in affected siblings are high.

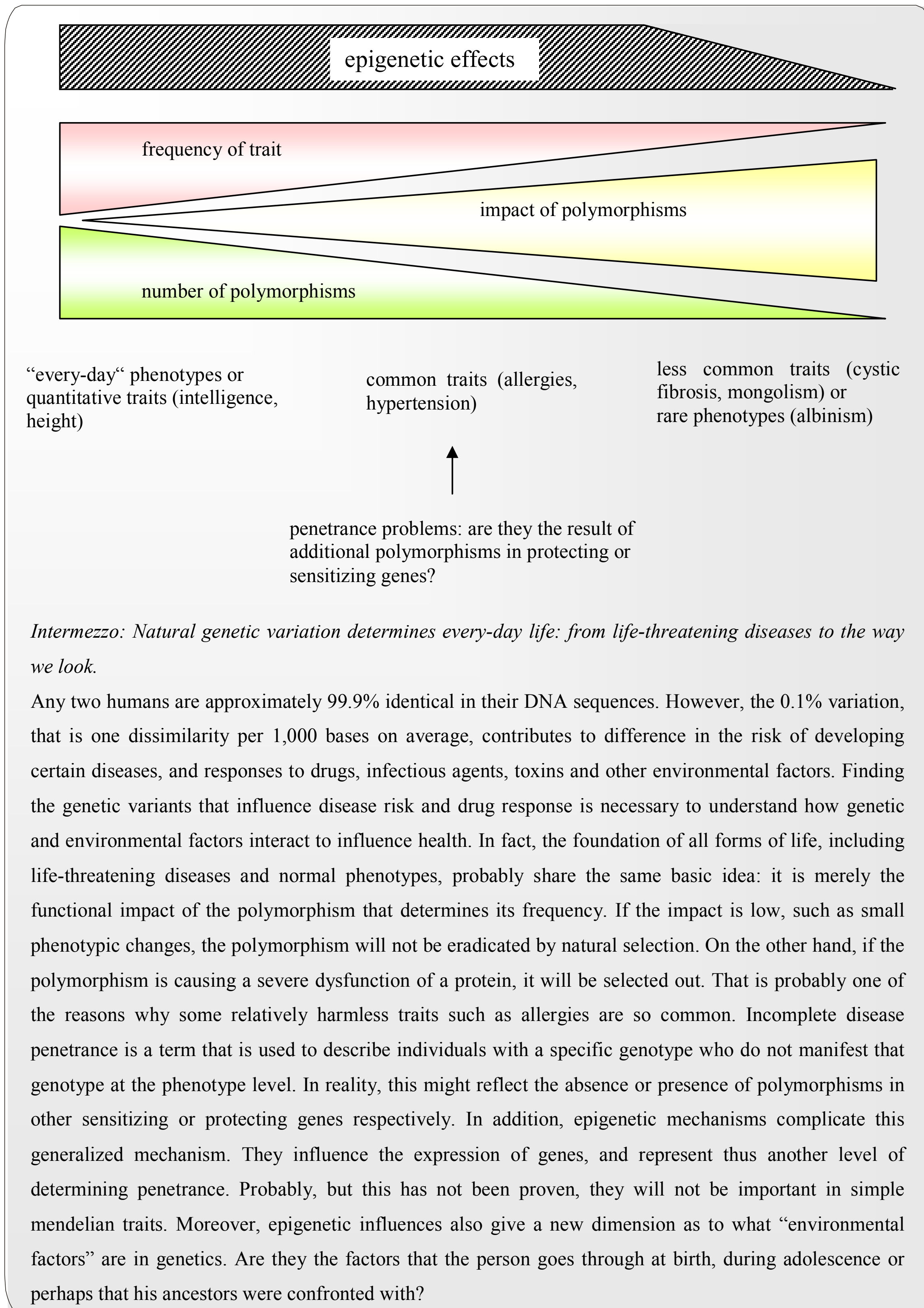
The principle treatment for CD is the lifetime controlling of gut inflammation, often resulting in surgical resection of parts of the intestine. However, because of substantial side effects and uncontrollable relapses, conventional treatment remains far from satisfactory for patients and physicians. A more rational approach for developing therapies should evolve from a good knowledge of the pathogenesis: when and why is inflammation initiated, how does it result in a chronic state and why is the patient not able to control the inflammation? A very important question is which genetic or environmental factors are common in most CD patients. It is generally accepted that good hygiene standards contributes significantly to the development of CD. This so-called hygiene-hypothesis is based on the reality that today, people encounter too little infections, resulting in a less trained immune system. Indeed, in developing countries, infectious diseases are fairly common, yet CD does practically not exist. Nevertheless, it is still a guess whether CD arises from the loss of tolerance to commensal bacteria or whether it is infection-triggered. In any case, the role of intestinal flora in genetically predisposed people is undisputable. Today, most studies are directed at this hypothesis.

In 2001, two independent research groups mapped the first susceptibility gene for CD, *CARD15*. This gene is located at the *IBD1* locus (16q12), a locus that was identified in almost every genome screen. The CARD15 protein is a cytosolic sensor for bacterial components. Up to one-third of CD patients have a probable loss-of-function mutation in this gene. The discovery of *CARD15* and CD-associated mutations was a major advance in unravelling some of the characteristics of CD pathogenesis. Firstly, it explains a great deal about the involvement of bacteria in the disease. For example, it was shown that epithelial cells overexpressing a mutated

CARD15 protein were not able to respond to infection with *Salmonella*. Secondly, these mutations are associated with ileal involvement. Interestingly, CARD15 is expressed in the paneth cells, which are specialized cells that secrete anti-bacterial molecules into the lumen of the gut. These paneth cells are normally restricted to the ileum, which fits with the phenotype of ileal affection in CD.

However, maybe *CARD15* should not be considered as being mutated in the strict meaning of the word (see also *Intermezzo*). In up to 20% of healthy persons, these mutations are also present. A gene like *CARD15* is most likely not a causative gene for CD, but rather a disease-modifying gene. Therefore, we better consider them as polymorphisms. An interesting hypothesis stated by Hugot and co-workers fits well within this concept. They speculated that these polymorphisms, which arose quite recently, were somehow beneficial during the outbreak of the plague in Europe. Later, the number of CD patients increased, partly because of increased hygiene standards. Because CARD15 is predominantly expressed in the ileum, patients carrying the "altered" CARD15 protein show more inflammation of the ileum.

The example of *CARD15* and the implications it had on research is substantial: a pubmed search on CARD15 gives 501 hits since 2001. Thus, the search for new functionally relevant proteins in the form of susceptibility genes is essential for a thorough understanding of mucosal biology in general and pathogenesis of CD in particular. The classical genetic linkage studies on CD have been crucial in the identification of *CARD15*. Therefore, we reviewed all the data of genetic linkage and gene association studies in **Chapter 1**. As for most complex genetic traits, many so-called chromosomal disease loci have been identified. Unfortunately, the exact genes within these loci that are associated with CD are not known. In contrast to simple mendelian traits, finding a master gene for the causation of CD is not plausible. In reality, the combination of many subtle polymorphisms with environmental risk factors and epigenetic mechanisms will most likely be responsible for disease onset. Another method for finding susceptibility genes is the candidate gene approach. Here, a gene is chosen based on a known property. Subsequently it is screened for mutations in several CD patients, and frequencies are compared in a larger case-control study. Though many genes have been screened in this manner, the reproducibility of association studies is disappointing. Mainly, this is the result of statistic setbacks and population biases. We gave an overview of all genes that were tested in such association studies. Two main conclusions can be drawn from this review. Firstly, patient and control cohorts should be large and thoroughly chosen. Often, populations are too small, and control cohorts are not age and sex matched. Good collaboration and communication of research labs with clinicians and hospitals is crucial. For most small laboratories, unfortunately, the use of undersized populations for



Intermezzo: Natural genetic variation determines every-day life: from life-threatening diseases to the way we look.

Any two humans are approximately 99.9% identical in their DNA sequences. However, the 0.1% variation, that is one dissimilarity per 1,000 bases on average, contributes to difference in the risk of developing certain diseases, and responses to drugs, infectious agents, toxins and other environmental factors. Finding the genetic variants that influence disease risk and drug response is necessary to understand how genetic and environmental factors interact to influence health. In fact, the foundation of all forms of life, including life-threatening diseases and normal phenotypes, probably share the same basic idea: it is merely the functional impact of the polymorphism that determines its frequency. If the impact is low, such as small phenotypic changes, the polymorphism will not be eradicated by natural selection. On the other hand, if the polymorphism is causing a severe dysfunction of a protein, it will be selected out. That is probably one of the reasons why some relatively harmless traits such as allergies are so common. Incomplete disease penetrance is a term that is used to describe individuals with a specific genotype who do not manifest that genotype at the phenotype level. In reality, this might reflect the absence or presence of polymorphisms in other sensitizing or protecting genes respectively. In addition, epigenetic mechanisms complicate this generalized mechanism. They influence the expression of genes, and represent thus another level of determining penetrance. Probably, but this has not been proven, they will not be important in simple mendelian traits. Moreover, epigenetic influences also give a new dimension as to what "environmental factors" are in genetics. Are they the factors that the person goes through at birth, during adolescence or perhaps that his ancestors were confronted with?

association studies is the only means of competing with the more organized labs. Nevertheless, even small scale pooling of cohorts can improve the outcome of genetic screenings. In 1997, a group of scientists and clinicians founded the IBD International Genetic Consortium (IBDIGC), now a group of twelve research groups involved in studying genes that are implicated in IBD. They collaboratively study large numbers of well-documented families for linkage. The formation of large international consortia offers research groups both large and small to participate equally in CD gene identification. The second conclusion is that more standardized methodologies for the screening of genes are required. In addition, we should evolve from single polymorphism association studies to the association of so-called haplotype structures. These haplotype blocks are the minimal genomic fragments that are genetically transmitted. A single polymorphism within a gene might not necessarily influence its function or transcription. Yet this defect could be caused by a nearby polymorphism within the same haplotype block. For instance, the *CARD15* gene contains, next to the CD-associated mutations, many other polymorphisms dispersed all over the gene. Because not all linkage observed at 16q12 can be explained by the three CD-associated polymorphisms, the question arises whether the actual disease causing mutation is located near the *CARD15* gene. In **Chapter 2**, we conducted a search for polymorphisms in the *CARD15* promoter region, because the expression of CARD15 is up-regulated during inflammation. We found a common polymorphism, c.-59G>A, located within a DNA stretch that potentially binds the E2F transcription factor. Using an *in vitro* reporter system, we showed that the promoter is less responsive to tumor necrosis factor alpha induction in the presence of the c.-59A allele. This functional defect prompted us to study this polymorphism in more detail. It appears that the -59A allele is always linked to the three other CD-associated mutations. This allele occurs isolated, but it is always present whenever one of the three CD-associated mutations occurs. Therefore, we decided to study the haplotype structure surrounding *CARD15*, which is currently ongoing.

The selection for a gene to study in a candidate gene approach might require some adjustment. The choice of a gene is very subjective, and virtually every gene can theoretically fit somewhere within the complexity of CD pathogenesis. We used a complementary and hypothesis-independent approach to identify more reliable candidate genes for CD in **Chapter 3**. A gene expression survey on unaffected colon biopsies from CD patients and healthy controls was used to identify genes that are differentially expressed in CD. We then focused on those genes that are located within an arbitrary region of 5 centimorgan from a marker that showed highest linkage to CD in at least two independent studies and in a study performed on a Flemish cohort. Eighteen genes were selected in this way. These genes can individually serve as new

focus genes, also for other researchers within the field. We focused on two genes within this list that belong to a cluster of highly related genes called metallothioneins (MT). These genes are located in tandem within the *IBD1* locus. Moreover, its role in host protection against various forms of stress and inflammation made them ideal candidate genes to pursue. We showed that their expression is significantly down-regulated in colon, ileum and blood of CD patients that have colonic disease. In addition, a reduced protein expression in this type of patients was shown by immunohistochemistry. Remarkably, we showed a correlation between RNA and protein levels, even though biopsies used for RNA extraction and for staining were not taken at the same moment. This further supports the idea that the altered MT expression is genetically determined. We found no coding or promoter mutations in one of the MT genes, *MT1M*. Therefore, we screened a well-characterized MT transcription factor, MRE-binding transcription factor 1, *MTF1*, for mutations. Two missense mutations were found in 6 out of 95 patients. Their functional relevance still needs to be determined. A frequent polymorphism in the splice site junctions at the boundary of exon 8 was found, however, we could not detect any splice variants. In our study, we focused on a polymorphism in the first intron, IVS-128A>T, which is a potential binding site for the GATA4 transcription factor. Although we showed that the frequency of this polymorphism was not different in CD patients as compared to healthy controls, it was significantly associated with disease location. The apparent resemblance to the association of *CARD15* with disease location encouraged us to study the combined risk for developing ileal disease. The odds for developing ileal disease was 4 in *CARD15* mutant patients, and 2.4 in patients that carry the IVS-128T allele. Moreover, we showed a gene-gene interaction of *CARD15* and *MTF1* as shown by a higher combined risk. Theoretically, this means that when CD patients are genotyped for *CARD15* and *MTF1*, ileal involvement (or alternatively, pure colonic disease) during disease progression can be predicted with substantial precision. Moreover, this is to our knowledge the first report of the use of microarray screening of tissues in genetic studies. We described the importance of gene expression profiling for the mapping of new susceptibility genes in the **Future prospects**, and provided a novel method based on transmission of expression signatures in family studies.

The involvement of MTF1 in the genetic down-regulation of MTs in CD patients needs to be determined. Interestingly, both CARD15 and MT are highly expressed in the ileum, and are involved in bacterial detection and eradication respectively. Bacterial load in the ileum is 10^{2-3} per gram faeces, while it reaches 10^{9-12} in the colon. It thus seems likely that the ileum is specialized in eliminating bacteria from its environment. CARD15 and MT may belong among the players that regulate this elimination. Conversely, the observation that mutant *CARD15* and

the IVS-128T allele predispose to ileal disease could even so be reversed. Possibly, a CARD15 mutated protein or the MTF1 protein in IVS-128T allele carriers, protects against an unwanted colonic inflammation, which might occur as a result of the high bacterial load at the colonic site. Importantly, the next priority is examining MT expression and assessing the risk of *MTF1* and *CARD15* in ulcerative colitis (UC) development, which is another related inflammatory bowel disease. Indeed, UC pathology could be considered as the extreme end of the disease location phenotype of CD. In UC patients, only the colon is affected but inflammation is disturbing only the superficial layers of the bowel wall. Although UC and CD are fundamentally different diseases, the concept of disease location might be influenced by the same molecular mechanisms.

We and others have demonstrated that MTs are highly expressed in intestinal epithelial cells (IEC). Over the past decade, many studies have revealed the physical and immunological importance of IEC in the gut homeostasis. These cells have developed a variety of mechanisms to reduce the risk of infection or damage by toxic compounds or free radicals. The effective maintenance of a physical barrier function is dependent on the establishment of well-organised intercellular junctions and a constant state of regeneration/renewal of the epithelium. Intestinal epithelial cells also participate in the innate immune responsiveness of the intestine by their ability to secrete mucus, antimicrobial peptides, cytokines and as well as chemokines. We generated an MT-knockdown model of HT29 colonic epithelial cells to study the biological consequences of a reduced MT expression in IEC. Initial experiments revealed that in comparison to the wild type cell line, these cells produce less IL8 when they were challenged invasive *E. coli*. This might provide a direct pathogenic role, since CD was previously linked to immunodeficiency caused by impaired innate immunity. In addition, preliminary data showed that these cells do not induce IL8 secretion in response to commensal bacteria (*Lactobacillus rhamnosus*), therefore an infection model for CD may fit best with our hypothesis. Nevertheless, a number of other bacterial strains, both commensal and pathogenic, need to be tested in this system.

Metallothioneins have been linked to a number of diseases, including some neurological disorders and different types of cancers. However, the regulation of MT isoforms in human cells is very complex, and its expression in health and disease is difficult to interpret. An overview of MT involvement in diseases, together with the regulation of MT expression is given in **Chapter 4**. For CD, conflicting data on MT expression are found in literature. The same holds true for MT expression in many cancers. The main problem in comparing these studies is the use of different techniques used for quantification. Since the original cloning of the first MT gene from

the renal cortex in 1960, original data on MT quantification in serum or on histological sections greatly depend on the antibody used. It is obvious that cross-reactivity between species but also between isoforms complicates the interpretation of results. These days, many researchers use the same anti-MT antibody, which facilitates comparing results. Recently, more detailed information on the expression on the different MT isoforms is gathered because of the emergence of more sensitive and specific techniques such as quantitative real-time PCR. In order to learn why MTs are not adequately expressed in CD, we made a literature overview of human MT regulation. This overview was based on an *in silico* comparison with murine expression control, because the majority of studies on MT were performed in mice. The genes upstream of MT are good candidates to study with respect to MT expression in CD.

In the next section, we used gene expression analysis to study the well-defined association between CD and another group of chronic inflammatory diseases called spondyloarthropathies (SpA). In addition, the genetic contribution of CD-associated *CARD15* mutations was assessed in these patients. SpA covers a heterogeneous group of arthritic diseases sharing clinical and radiological features and presence of the *HLA-B27* allele. Previous research has indicated that patients with spondyloarthropathy (SpA) form a population with a substantial higher risk for the development of CD, which is discussed in **Chapter 5**. Subclinical inflammatory gut lesions are common in SpA, as observed during colonoscopic studies in these patients. These gut lesions are either classified as acute or chronic inflammation. The latter form resembles closely lesions found in CD patients. Follow-up studies of SpA patients indicated that 7% develop an inflammatory bowel disease and, of those with chronic inflammatory gut lesions, 30% develop clinical CD. Several lines of evidence suggest that the gut could have an important pathogenic role in SpA. However, the precise relation of the join-gut axis is not completely understood. We therefore assessed the contribution of *CARD15* mutations in SpA patients, as described in **Chapter 6**. We found that the frequency of *CARD15* mutations in SpA patients with chronic gut lesions was significantly higher than in the control population, and comparable to the frequency in CD patients. Thus, we provided a first line of evidence that gut inflammation in SpA patients is genetically determined, and that SpA patients with subclinical chronic gut inflammation are a separate group, with a higher risk to develop CD.

We wondered whether the intestinal gene expression pattern in SpA patients differs from that of controls. In **Chapter 7** we described alterations in gene expression in unaffected colon biopsies of SpA patients as compared to controls. Furthermore, the alterations found in patients with a history of chronic gut inflammation are comparable to those found in CD patients,

confirming the clinical observations. Hence, SpA serves as a good model to study early CD. Furthermore, the SpA model might be used in the selection for potential candidate genes for the predisposition of gut inflammation.

Samenvatting en discussie

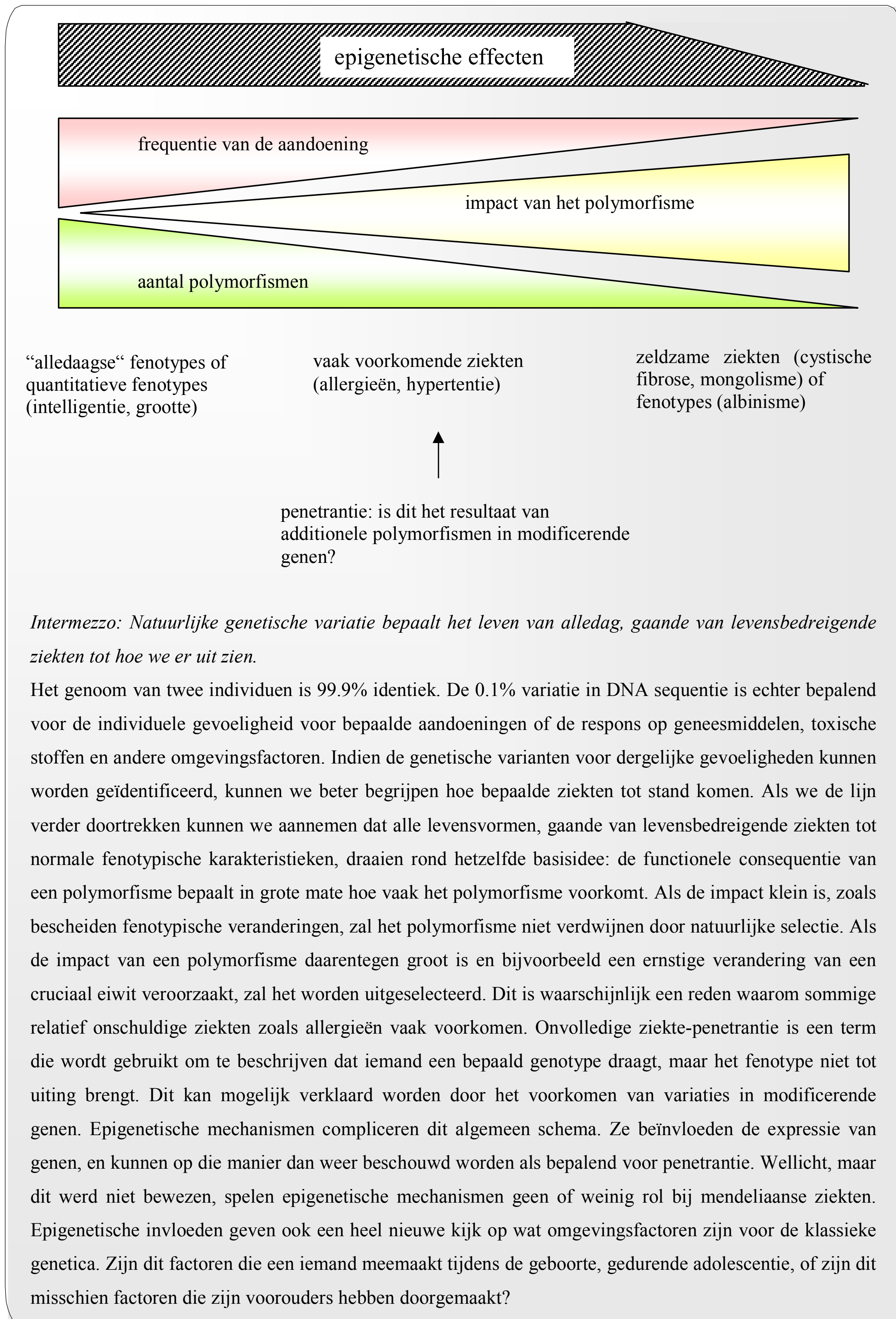
De opzet van dit project bestond erin genexpressie te bestuderen om genetische afwijkingen voor de ziekte van Crohn (ZVC) op te sporen en te karakteriseren. Patiënten met de ZVC lijden aan een chronische ontsteking in de darmen. Hierbij is de volledige darmwand, zowel de oppervlakkige als de diepere lagen, aangetast. Het ziektebeeld is sterk individueel bepaald en kent een grillig verloop: periodes van hevige klachten en kalmere periodes wisselen elkaar af. De belangrijkste verschijnselen van de ZVC zijn een rechtstreeks gevolg van de ontsteking zelf, zoals pijn in de buik, koorts, krampen, al dan niet bloederige diarree en gewichtsverlies. Vooral het uiteinde van de dikke darm (colon) en de dunne darm (ileum), maar ook andere delen van het maag-darmstelsel kunnen worden aangetast. Het is zelfs zo dat de plaats van ontsteking wordt gebruikt om patiënten met de ZVC te classificeren, omdat deze relatief stabiel blijft tijdens het ziekteverloop. Bovendien is er een sterke overeenkomst in de plaats van inflammatie tussen verwanten met de ZVC.

De ZVC kan niet worden genezen, dus is de behandeling erop gericht de ontsteking onder controle te houden en complicaties te verminderen. Vaak moeten patiënten na verloop van tijd een chirurgische ingreep ondergaan, gaande van het weghalen van een abces tot het verwijderen van een deel van de zieke darm of zelfs het volledige colon. Hoewel veel patiënten aanvankelijk geholpen zijn met de standaardbehandelingen treden er vaak neveneffecten op, en kunnen perioden van herval niet worden gecontroleerd. Een meer rationele aanpak gebaseerd op kennis van het ziekteproces is noodzakelijk bij het ontwikkelen van nieuwe behandelingen. Hierbij moeten we enkele concepten proberen te begrijpen: wanneer en hoe wordt de ziekte geïnitieerd, waarom evolueert een ontsteking tot chronisch lijden, en waarom is de patiënt niet in staat de inflammatie onder controle te houden? Belangrijk is ook de vraag welke genetische en omgevinsfactoren patiënten gemeenschappelijk hebben. Een algemeen aanvaarde hypothese is de zogenaamde hygiëne-hypothese. Ze gaat ervan uit dat we tegenwoordig te weinig infectieziekten doormaken, waardoor het afweersysteem onvoldoende geoefend wordt. Inderdaad, het is opmerkelijk dat er bijna geen ZVC voorkomt in ontwikkelingslanden, waar er echter veel infectiegevaar dreigt. Nochtans weten we nog altijd niet of patiënten die de ziekte ontwikkelen een verminderde tolerantie vertonen in de darm, of dat de ziekte een initiële infectie vereist. Hoe dan ook, we gaan er van uit dat de darmflora een belangrijke rol speelt in de ZVC, maar enkel bij personen met een genetische aanleg de ziekte uitlokt. Dit is dan ook vandaag de dag het onderwerp van veel onderzoek in de ZVC.

Het eerste gen dat sterk werd geassocieerd met de ZVC werd gekloneerd in 2001. Dit gen, *CARD15* genaamd, ligt in de *IBD1* locus (16q12), een locus die in bijna elke genoomscan werd geïdentificeerd. Het CARD15 eiwit herkent bacteriële componenten binnen in de cel, en activeert NFκB. Ongeveer een derde van de patiënten met de ZVC dragen een mutatie in dit gen. Veel onderzoek was aanvankelijk gebaseerd op de functie van deze mutaties, maar hierover bestaat nog geen sluitend antwoord. Er zijn echter sterke aanwijzingen dat het gemuteerde CARD15 eiwit zijn functie gedeeltelijk niet meer kan uitoefenen. Over het algemeen was de ontdekking van *CARD15* van groot belang om enkele aspecten in het ziekteverloop van de ZVC te begrijpen. Ten eerste bieden ze een verklaring voor het belang van bacteriën in het ziekteproces. Een mooi voorbeeld hiervan is dat epitheliale cellen die het gemuteerde CARD15 eiwit overexpresseren niet meer reageren op infectie met *Salmonella*. Ten tweede, mutaties in *CARD15* werden gecorreleerd met ileale aantasting van de ZVC, de plaats waar CARD15 sterk tot expressie komt. CARD15 wordt er geëxpresseerd in paneth cellen, gespecialiseerde cellen die bacteriën in het ileum elimineren. In normale omstandigheden komen deze cellen enkel voor in het ileum, wat een directe link vormt tussen de werking van CARD15 en de plaats van aantasting.

Misschien is het belangrijk om de mutaties in *CARD15* eerder te beschouwen als polymorfismen die een invloed uitoefenen op het ziekteproces (zie ook *Intermezzo*). Ze komen ook voor bij 20% van gezonde personen. Het is geen oorzakelijk gen, maar eerder een gen dat een invloed heeft op de manifestatie van de ziekte, in dit geval, ileale aantasting. Deze redenering sluit goed aan bij een hypothese die Hugot en medewerkers vooropstelden. Zij geloven dat personen die deze polymorfismen in *CARD15* droegen, bevoordeeld waren bij het uitbreken van de pest in Europa. De oorsprong van deze polymorfismen loopt inderdaad gelijk met deze periode. Vanaf dan is de incidentie van de ZVC beginnen stijgen, mede door een hogere hygiënestandaard. Personen met mutaties in *CARD15* en de ZVC ontwikkelen voornamelijk ontsteking in het ileum, waar CARD15 voornamelijk tot expressie komt.

De impact van *CARD15* op het onderzoek naar de ZVC, maar ook naar aangeboren immuniteit is groot: het resultaat voor CARD15 in pubmed geeft 501 hits sinds 2001. Dit betekent dat de identificatie van nieuwe eiwitten betrokken bij genetische predispositie, van belang is bij de studie van mucosale biologie en van de ZVC in het bijzonder. De klassieke genetische *linkage* studies waren cruciaal bij het kloneren van *CARD15*. Daarom hebben we een overzicht gemaakt van alle *linkage* en associatie studies bij de ZVC in ***Chapter* 1**. Zoals voor de meeste complex genetische aandoeningen werden er meerdere chromosomale loci gevonden voor de ZVC. De precieze genen binnen die loci die geassocieerd zijn met ZVC werden echter



Intermezzo: Natuurlijke genetische variatie bepaalt het leven van alledag, gaande van levensbedreigende ziekten tot hoe we er uit zien.

Het genoom van twee individuen is 99.9% identiek. De 0.1% variatie in DNA sequentie is echter bepalend voor de individuele gevoeligheid voor bepaalde aandoeningen of de respons op geneesmiddelen, toxische stoffen en andere omgevingsfactoren. Indien de genetische varianten voor dergelijke gevoeligheden kunnen worden geïdentificeerd, kunnen we beter begrijpen hoe bepaalde ziekten tot stand komen. Als we de lijn verder doortrekken kunnen we aannemen dat alle levensvormen, gaande van levensbedreigende ziekten tot normale fenotypische karakteristieken, draaien rond hetzelfde basisidee: de functionele consequentie van een polymorfisme bepaalt in grote mate hoe vaak het polymorfisme voorkomt. Als de impact klein is, zoals bescheiden fenotypische veranderingen, zal het polymorfisme niet verdwijnen door natuurlijke selectie. Als de impact van een polymorfisme daarentegen groot is en bijvoorbeeld een ernstige verandering van een cruciaal eiwit veroorzaakt, zal het worden uitgeselecteerd. Dit is waarschijnlijk een reden waarom sommige relatief onschuldige ziekten zoals allergieën vaak voorkomen. Onvolledige ziekte-penetrantie is een term die wordt gebruikt om te beschrijven dat iemand een bepaald genotype draagt, maar het fenotype niet tot uiting brengt. Dit kan mogelijk verklaard worden door het voorkomen van variaties in modificerende genen. Epigenetische mechanismen compliceren dit algemeen schema. Ze beïnvloeden de expressie van genen, en kunnen op die manier dan weer beschouwd worden als bepalend voor penetrantie. Wellicht, maar dit werd niet bewezen, spelen epigenetische mechanismen geen of weinig rol bij mendeliaanse ziekten. Epigenetische invloeden geven ook een heel nieuwe kijk op wat omgevingsfactoren zijn voor de klassieke genetica. Zijn dit factoren die een iemand meemaakt tijdens de geboorte, gedurende adolescentie, of zijn dit misschien factoren die zijn voorouders hebben doorgemaakt?

nog niet geïdentificeerd. In tegenstelling tot mendeliaanse genetische ziekten verwacht men geen geïsoleerd dominant gen dat de oorzaak is bij alle patiënten met de ZVC. Wellicht zal de combinatie van vele polymorfismen, omgevingsfactoren en epigenetische mechanismen aan de basis liggen van de ZVC. Een veel gebruikte alternatieve manier om genen op te sporen is de zogenaamde kandidaatgen benadering. Hierbij wordt er een arbitrair gen gekozen dat mogelijk betrokken is in de ziekte, en zoekt men bij een klein aantal patiënten naar mutaties in dat gen. Vervolgens wordt de frequentie van eventuele mutaties vergeleken bij een groot aantal patiënten en controles. Veel genen en mutaties werden op deze manier geassocieerd met de ZVC, maar de reproduceerbaarheid is telkens minimaal. Dit is voornamelijk het gevolg van beperkingen in statistische methoden en populatie *bias*. In *Chapter* 1 geven we een overzicht van alle genen die werden getest op associatie. Hieruit kunnen we twee belangrijke conclusies trekken. Ten eerste moet de keuze van de testpopulaties goed overwogen worden, en moet deze voldoende groot zijn. Vaak zijn de populaties te klein, en zijn controles niet gelinkt aan leeftijd en geslacht met de patiëntenpopulatie. Een goede samenwerking tussen onderzoekslaboratoria, artsen en hospitalen is cruciaal. Zelfs het poolen van cohorten op kleine schaal kan grote impact hebben op de uitkomst van genetische studies. In 1997 stichtten enkele wetenschappers en artsen het *IBD International Genetic Consortium* (IBDIGC). Vandaag telt die groep twaalf onderzoeksgroepen die werken rond genetica van inflammatoire darmaandoeningen. Dit consortium biedt ook kleine groepen de mogelijkheid om op gelijke basis deel te nemen aan de identificatie van genen voor de ZVC. De tweede conclusie die kan worden getrokken uit de vele associatiestudies is dat er meer gestandaardiseerde methodologieën moeten gebruikt worden bij het screenen van genen. Daarenboven zouden we moeten evolueren van de analyse van geïsoleerde polymorfismen naar zogenaamde haplotype blokken. Deze haplotype structuren zijn de minimale DNA fragmenten die worden overgeërfd. Het is best mogelijk dat een geïsoleerd polymorfisme in een gen geen effect heeft op zijn expressie of de functie van zijn eiwit, maar dat een mutatie verderop, binnen hetzelfde haplotype blok, wel een effect heeft. Een mooi voorbeeld hiervan is dat er binnen het *CARD15* gen vele polymorfismen werden geïdentificeerd, waarvan er drie werden geassocieerd met de ZVC. Men kan zich nu afvragen of een andere mutatie binnen het haplotype blok dat *CARD15* bevat, verantwoordelijk is voor een gewijzigde functie van CARD15. In ***Chapter* 2** hebben we gezocht naar polymorfismen in de promoter regio van *CARD15*, mede omdat de expressie van CARD15 opgereguleerd wordt bij inflammatie. We beschrijven een polymorfisme, c.-59G>A, in een regio dat mogelijk de E2F transcriptiefactor bindt. In een *in vitro* reportersysteem toonden we aan dat het c.-59A allel minder responsief was op *tumor necrosis factor alpha* inductie. Dit heeft ons ertoe aangezet om dit polymorfisme meer in detail te

bestuderen. Dit polymorfisme komt altijd samen voor met de andere mutaties in *CARD15*, maar het kan ook geïsoleerd voorkomen. Daarom zijn we gestart met het karakteriseren van het haplotype blok dat *CARD15* omvat, om na te gaan of het c.-59G>A polymorfisme deel uitmaakt van een haplotype blok dat geassocieerd kan worden met de ZVC. Dit werk is nu gaande.

Er is nood aan additionele aanwijzingen bij de keuze van genen om te testen in associatie studies. Inderdaad, de keuze van een gen is vaak subjectief, en elk gen kan bij wijze van spreken functioneel gekoppeld worden aan de ZVC. Daarom gebruikten we in ***Chapter* 3** een alternatieve en hypothese-onafhankelijke aanpak om nieuwe en meer betrouwbare kandidaatgenen te identificeren voor de ZVC. In eerste instantie werd een genexpressie-analyse uitgevoerd in gezonde colon biopten. Er werden differentiële geëxpresseerde genen geïdentificeerd tussen patiënten met de ZVC en gezonde controles. Vervolgens selecteerden we genen die binnen een range van 5 centimorgan rond een merker lagen. De merkers die we hiervoor gebruikten lagen in een locus die in ten minste twee onafhankelijke studies werd geïdentificeerd, of in loci die werden geidentificeerd in een studie uitgevoerd met een vlaamse cohorte. Op deze manier selecteerden we 18 genen. Elk van deze genen kunnen dienen als focus genen voor eenieder die geïnteresseerd is in genetica van de ZVC. Wij hebben ons toegelegd op het karakteriseren van twee genen binnen deze lijst van 18, de metallothioneines (MT). Deze twee genen behoren tot een grote genfamilie die allen in tandem gelegen zijn binnen de *IBD1* locus. Een tweede reden om deze genen te bestuderen is dat ze een belangrijke rol spelen bij de verdediging van het organisme tegen verschillende vormen van stress en inflammatie. We toonden aan dat deze genen minder tot expressie komen in het colon, ileum en bloed van patiënten met de ZVC met colon aantasting. Dit werd bevestigd op eiwitniveau. Een interessante bevinding was dat het RNA- en eiwitniveau gecorreleerd waren binnen dezelfde patiënten, hoewel het weefsel voor beide technieken niet op dezelfde dag werd genomen. Dit en de algemeen verminderde expressie in bloed toonde aan dat er een mogelijk genetisch defect is bij de expressie van MT in patiënten met de ZVC die colon aantasting hebben. We vonden geen mutaties in coderende sequenties of in de promoter regio van één van de MT genen, *MT1M*. Omdat de verminderde expressie van meerdere MT isovormen werd aangetoond, besloten we om de best gekarakteriseerde transcriptiefactor van MT, *MRE-binding transcription factor 1 of MTF1*, te screenen op mutaties. We vonden twee mutaties die een aminozuurverandering teweeg brachten in 6 van de 95 patiënten met de ZVC. De functionele relevantie hiervan moet echter nog worden bestudeerd. Verder beschreven we een polymorfisme in een *splice acceptor* positie, maar we hebben geen alternatieve *splice* variant kunnen aantonen. Wij hebben ons toegelegd op een polymorfimse in het eerste intron van *MTF1*, IVS-128A>T, omdat deze regio mogelijk de transcriptiefactor

GATA4 kan binden. Er was geen verschil in frequentie van dit polymorfisme in controles en patiënten, maar het heeft net zoals *CARD15* een grote invloed op de locatie van de ziekte. Het risico om ileale aantasting te krijgen is bijzonder hoog als het IVS-128T allel en een van de *CARD15* mutaties aanwezig zijn: de odds ratio is 4 in patiënten met een *CARD15* mutatie, en 2.4 in patiënten met het IVS-128T allel. We toonden ook aan dat er een gen-gen interactie is tussen *CARD15* en *MTF1*, aangezien er een hoger gecombineerd risico is dan verwacht in aanwezigheid van beide factoren. In theorie kan dit betekenen dat we door genotypering van patiënten met de ZVC goed kunnen voorspellen of ze in de loop van het ziekteproces ileale inflammatie, of zuivere colon aantasting zullen ontwikkelen. Tot slot, deze studie is de eerste waar genexpressie wordt gebruikt bij genetische studies. We benadrukken het belang van het bestuderen van genexpressie in predispositie voor complexe ziekten in de ***Future prospects***, en beschrijven een nieuwe methode om kandidaatgenen te identificeren aan de hand van de overerving van expressiepatronen binnen families.

De manier waarop MTF1 een invloed kan hebben op de expressie van MT in de ZVC vereist nadere studie. Zowel CARD15 als MT worden voornamelijk geëxpresseerd in het ileum, en zijn betrokken bij respectievelijk de herkenning en eliminatie van bacteriën. Er is een groot verschil in aantal en soort bacteriën tussen het ileum en het colon: in het ileum bedraagt het aantal bacteriën ongeveer 10^{2-3} per gram faeces, terwijl dat in het colon oploopt tot 10^{9-12}. Men kan dus stellen dat het ileum een gespecialiseerde omgeving is waar het aantal bacteriën gelimiteerd moet worden. CARD15 en MT zouden in dit proces een rol kunnen spelen. De observatie dat *CARD15* mutaties en het IVS-128T allel geassocieerd zijn met ileale aantasting kan ook anders worden geïnterpreteerd: *CARD15* mutanten en dragers van het IVS-128T allel kunnen beschermd zijn tegen zuivere colon aantasting, omdat de mogelijkheid om een inflammatoire respons op te wekken verminderd is. De hoogste prioriteit is nu om de MT expressie en het IVS-128A>T polymorfisme in combinatie met *CARD15* mutaties te bepalen in patiënten met *colitis ulcerosa* (CU), een andere vorm van inflammatoir darmlijden. Deze ziekte kan tot op zekere hoogte beschouwd worden als het extreme uiteinde van een continu ziektebeeld van ZVC. Bij patiënten met CU is enkel het colon ontstoken, maar alleen de oppervlakkige lagen van de darmwand zijn aangetast. Hoewel CU en de ZVC fundamenteel verschillende ziekten zijn, is het niet ondenkbaar dat er gemeenschappelijke erfelijke factoren bestaan die een rol spelen bij het tot stand komen van inflammatie op een specifieke plaats in de darm.

We toonden aan dat MTs voornamelijk geëxpresseerd worden in epitheliale cellen, een eigenschap die al eerder werd beschreven. De laatste jaren werd het belang van intestinale

epitheelcellen in immunologie aangetoond. Deze cellen bevatten een aantal mechanismen om de kans op bacteriële infectie of schade door toxische stoffen te reduceren. Ze zijn ook betrokken in aangeboren immuniteit: ze produceren anti-bacteriële stoffen en mucus, maar ook cytokines en chemokines. We hebben HT29 colon-epitheelcellen gegenereerd die MT verminderd tot expressie brengen, d.m.v. *small interfering RNA*. We toonden aan dat deze cellen een verminderde IL8-secretie vertonen in vergelijking met de normale HT29 cellen wanneer ze geïnfecteerd werden met invasieve *E. coli*. Deze observatie is een eerste directe link tussen een verlaagde MT expressie en de ZVC: deze wordt immers gekarakteriseerd door immunodeficiëntie als gevolg van een gestoorde *innate* immuniteit. Preliminaire resultaten toonden aan dat deze cellen niet reageren op commensale bacteriën (*Lactobacillus rhamnosus*). Dit betekent dat een infectiemodel eerder aanleunt bij onze bevindingen. Om hierover zekerheid te krijgen moeten we echter een reeks bacteriestammen testen, zowel pathogene als niet-pathogene stammen.

Metallothioneines werden gelinkt aan een aantal ziekten, waaronder enkele neurologische aandoeningen en kankers. De expressie van humane MT isovormen is echter zeer complex, en de interpretatie van de resultaten is vaak erg moeilijk en onderling niet te vergelijken. In ***Chapter* 4** geven we een overzicht van de expressie van MT in verschillende ziekten, alsook de regulatie van MT in humane cellen. Er bestaan een aantal tegenstrijdige resultaten in verband met MT en de ZVC. Hetzelfde probleem geldt voor MT en kankers. De voornaamste reden hiervoor is dat er verschillende technieken werden gebruikt voor het kwantificeren van MT. Sinds het eerste MT gen werd gecloneerd in 1960, verschenen er veel studies die MT maten in serum en op histologie aan de hand van antilichamen. Cross-reactiviteit tussen species, maar ook tussen verschillende isovormen maakt het moeilijk om deze data te vergelijken. Vandaag wordt echter meestal hetzelfde anti-MT antilichaam gebruikt. Daarenboven maken meer gespecialiseerde technieken, zoals kwantitatieve PCR, het mogelijk om gedetaileerde informatie te krijgen over de expressie van verschillende isovormen. Om een breed beeld te krijgen over hoe een verlaagde MT expressie kan verklaard worden in de ZVC maakten we een overzicht van de regulatie van MT in humane systemen. Dit overzicht is gebaseerd op een *in silico* vergelijking tussen humane en muis genregulatie, aangezien de meerderheid van de studies gebeurd is op muizen. De genen die aan de basis liggen van MT regulatie zijn uiteraard kandidaatgenen om te testen in de ZVC.

In het tweede gedeelte van het project hebben we gebruik gemaakt van genexpressie om het verband te bestuderen tussen de ZVC en een andere chronisch inflammatoire aandoening,

spondyloartropathie (SpA). Daarenboven werden de *CARD15* mutaties bestudeerd in deze groep van patiënten. SpA is een heterogene groep van verwante aandoeningen gekarakteriseerd door specifieke radiologische en klinische aspecten. Ze vertonen ook een sterke associatie met het *HLA-B27* allel. Er zijn veel aanwijzingen, beschreven in ***Chapter* 5**, dat SpA patiënten een risicogroep vormen voor het ontwikkelen van de ZVC. Darmlaesies zijn een vaak voorkomend fenomeen bij SpA patiënten, zij het in subklinische vorm. Deze laesies kunnen acuut of chronisch van aard zijn. De chronische laesies lijken sterk op de laesies die gevonden worden bij patiënten met de ZVC. Een studie toonde aan dat 7% van de SpA patiënten de ZVC of CU ontwikkelen, en meer specifiek 30% van de SpA patiënten met chronische darminflammatie ontwikkelden de ZVC. Een aantal argumenten wijzen op het belang van deze subklinische darminflammatie bij het ontstaan van SpA. Hoe dan ook, het precieze verband tussen het gewricht en de darm is nog niet volledig begrepen. In ***Chapter* 6** beschrijven we een studie waarbij we de ZVC-geassocieerde *CARD15* mutaties in SpA patiënten beschouwen. We toonden aan dat de frequentie van *CARD15* mutaties in SpA patiënten met subklinische chronische darminflammatie was verhoogd in vergelijking met controles, en vergelijkbaar met de frequentie in patiënten met de ZVC. Met deze studie toonden we aan dat darminflammatie in SpA mede genetisch bepaald is. SpA patiënten met chronische darminflammatie zijn inderdaad te beschouwen als een aparte groep die een verhoogd risico vormen voor het ontwikkelen van de ZVC, mede door het dragen van *CARD15* mutaties.

We stelden ons de vraag of de darm bij SpA patiënten verschillen vertoont in genexpressie in vergelijking met controles. In ***Chapter* 7** beschrijven we afwijkingen in genexpressie van normale colon biopten bij SpA patiënten. Deze afwijkingen zijn bij patiënten met chronische darminflammatie te vergelijken met afwijkingen in de ZVC, wat de originele klinische data bevestigen. SpA kan daarom dienen als een pre-crohn model, ook bij het selecteren van kandidaatgenen voor darminflammatie.

Curriculum Vitae

Debby Laukens (°13/08/1977, Antwerpen, Belgium)

AFFILIATION

Department for Molecular Biomedical Research
Ghent University
Technologiepark 927, B-9052 Zwijnaarde
Tel: +329/ 331 36 64
Fax: +329/ 331 3609
E-mail: debby.laukens@dmbr.ugent.be

PRESENT AFFILIATION

Department of Gastroenterology
Ghent University Hospital
De Pintelaan 185, B-9000 Gent
Tel: +329/ 240 20 64
Fax: +329/ 240 49 84
E-mail: debby.laukens@ugent.be

EDUCATION AND ACADEMIC CAREER

1995-1997: Candidate (Bachelor) in Biology (University of Antwerp, RUCA)
1998-1999: Licentiate (Master) in Biochemistry (University of Antwerp, UA)

Master thesis: "Localization and characterization of ataxin-7 in the mouse" at the laboratory of Molecular Genetics, University of Antwerp.
Promoters: Claire Mauger and Christine Van Broeckhoven

1999-2000: PhD fellowship at the laboratory of Molecular Genetics, University of Antwerp and Flanders Interuniversity Institute for Biotechnology
Promoters: Claire Mauger and Christine Van Broeckhoven

2000-present: PhD fellowship at the Department for Molecular Biomedical Research, University of Ghent and the Flanders Interuniversity Institute for Biotechnology.
Promoters: Pieter Rottiers and Erik Remaut

SCIENTIFIC ACTIVITIES

Publications in international scientific journals

Debby Laukens, Denis Marichal, Marthe Van Den Berghe, Harald Peeters, Dirk Elewaut, Filip De Keyser, Herman Mielants, Claude Cuvelier, Eric M.Veys, Erik Remaut, Lothar Steidler and Martine De Vos (2005). "*CARD15* mutations in patients with spondyloarthropathy are linked with disease progression and evolution to Crohn's disease." Ann Rheum Dis. **64**(6): 930-935.

Harald Peeters, Bert Vander Cruyssen, **Debby Laukens**, Paul Coucke, Denis Marichal, Marthe Van Den Berghe, Claude Cuvelier, Erik Remaut, Herman Mielants, Filip De Keyser and Martine De Vos (2004). "Radiological sacroiliitis, a hallmark of spondylitis, is linked with *CARD15* gene polymorphisms in patients with Crohn's disease." Ann Rheum Dis. **63**: 1131-1134.

Bert Vander Cruyssen, Harald Peeters, Hoffman IE, **Debby Laukens**, Paul Coucke, Denis Marichal, Claude Cuvelier, Erik Remaut, Eric M Veys, Herman Mielants, Martine De Vos and Filip De Keyser (2005). "*CARD15* polymorphisms are associated with anti-Saccharomyces cerevisiae antibodies in caucasian Crohn's disease patients." Clin Exp Immunol. **140**: 354-359.

Harald Peeters, Sara Bogaert, **Debby Laukens**, Pieter Rottiers, Filip De Keyser, Arlette Darfeuille-Michaud, Dirk Elewaut and Martine De Vos (2006). "*CARD15* variants determine a disturbed response of monocytes to *Adherent-Invasive Escherichia Coli LF82* in Crohn's Disease." *Submitted to IBD*.

Debby Laukens, Harald Peeters, Bert Vander Cruyssen, Tom Boonefaes, Dirk Elewaut, Filip De Keyser, Herman Mielants, Claude Cuvelier, Eric Veys, Kenny Knecht, Paul Van Hummelen, Erik Remaut, Lothar Steidler, Martine De Vos, and Pieter Rottiers (2006). "Altered gut transcriptome in Spondyloarthropathy." Ann Rheum Dis. *Accepted.*

Debby Laukens, Harald Peeters, Sara Bogaert, Anouk Waeytens, Bert Vander Cruyssen, Dirk Elewaut, Filip De Keyser, Herman Mielants, Claude Cuvelier, Eric M. Veys, Kenny Knecht, Paul Van Hummelen, Jurgen Del-Favero, Erik Remaut, Martine De Vos and Pieter Rottiers (2006). "Reduced metallothionein expression in colonic Crohn's disease: evidence for *MTF1* as a new disease-modifying gene." *Submitted to Gastroenterology*.

Debby Laukens, Anouk Waeytens, Pieter De Bleser, Claude Cuvelier, Martine De Vos and Pieter Rottiers. (2006). "Human metallothionein expression under normal and pathological conditions: mechanisms of gene-regulation." *Review in preparation.*

Abstracts in international scientific journals

Martine De Vos, **Debby Laukens**, Denis Marichal, *et al.* (2003). "*CARD15* mutations in patients with spondyloarthropathy are linked with disease progression and evolution to Crohn's disease." Gastroenterology. 124 (suppl): A375.

Participation of international meetings

Symposium on auto-immunity, Antwerpen, Belgium, November 30, 2001
Participant

3[d] International congress on spondyloarthropathies, Gent, Belgium, October 2-5, 2002
Participant

The regulation of mucosal inflammation, Keystone, Colorado, USA, April 1-6, 2003
Oral presentation "*CARD15* mutations in patients with spondyloarthropathy are linked with disease progression and evolution to Crohn's disease."

Annual Meeting Society for Mucosal Immunology 2003, Paris, France, May 15, 2003
Participant

12th International congress of immunology, Montreal, Canada, July 19-22, 2004
Oral presentation "Transcriptome analysis of pre-crohn patients in a cohort of spondyloarthropathy patients."

4th International congress on spondyloarthropathy, Gent, Belgium, October 8, 2004
Poster presentation "Transcriptome analysis of pre-crohn patients in a cohort of spondyloarthropathy patients."

Inflammatory bowel disease, research drives clinics, Munster, Germany, September 2-3, 2005
Poster presentation "Identification of susceptibility genes for Crohn's disease via transcriptome analysis."

Digestive disease week 2006, Los Angeles, California, USA, May 20-25, 2006
Poster presentation "Identification of susceptibility genes for Crohn's disease via transcriptome analysis."

Participation of national meetings

3th VIB seminar, Mol, Belgium, February 17-18, 2000

4th VIB seminar, Blankenberge, Belgium, February, March 15-16, 2001

5th VIB seminar, Blankenberge, Belgium, March 08-09, 2002

6th VIB seminar, Blankenberge, Belgium, March 13-14, 2003
Poster presentation "*CARD15* mutations in patients with spondyloarthropathy are linked with disease progression and evolution to Crohn's disease."

Doctoraatssymposium, ICC, Gent, Belgium, April 30, 2003
Poster presentation "*CARD15* mutations in patients with spondyloarthropathy are linked with disease progression and evolution to Crohn's disease."

7th VIB seminar, Blankenberge, Belgium, March 11-12, 2004
Participant

4th VIB MicroArray user group meeting, Brussel, Belgium, November 19, 2004
Oral presentation "Identification of susceptibility genes for Crohn's disease via transcriptome analysis."

8th VIB seminar, Blankenberge, Belgium, March 3-4, 2005
Participant

9th VIB seminar, Blankenberge, Belgium, March 09-10, 2006
Oral presentation "Alternative method to identify susceptibility genes for Crohn's disease through transcriptome analysis."

Awards

NIH grant award (Department of Health and Human Services, National Institutes of Health, National Institute of Allergy and Infectious Disease: Grant Number 1R13 AI 053891-01) for attending the meeting entitled "The Regulation of Mucosal Inflammation", Keystone Resort, Colorado, USA.

FOCIS 2004 Travel Award (Federation of Clinical Immunology Societies), abstract submission entitled "Transcriptome Analysis of Pre-Crohn Patients in a Cohort of Spondyloarthropathy Patients." 12th International congress for Immunology, Montreal, Canada.

De zoveelste rit van Gent naar huis, onderweg is er dan altijd veel tijd om na te denken. Ik voel een emotionele bui aankomen - er staat dan ook vanalles te veranderen - en ik besluit om mijn dankwoord te schrijven van zodra ik thuiskom. Zonet had ik nog een fijn gesprek met **Sarah**, dat gebeurt wel vaker. Zij vertrekt binnenkort naar Madrid, ik wens haar het allerbeste en zal ze missen! Ik ben thuisgekomen. Lichten aan, verwarming aan, de poes die rond mijn benen draait en mij zoals gewoonlijk bijna doet struikelen. Chris, je bent er niet vanavond. Ik heb honger, het is al 20u, en ik heb geen zin om te koken. Een beetje soep opwarmen en een pizza in de oven. Ik zit ongeveer halverwege mijn doctoraat, en ik vraag mij af waarom ik het toch zo moeilijk heb om door te werken, en alles op tijd af te krijgen. Gisteren kreeg ik te horen van **Martine De Vos** dat ik moet leren om meer gestructureerd te zijn, en ze heeft overschot van gelijk. Binnenkort kan ik beginnen in haar groep, ik kijk er naar uit, vooral ook omdat ik dan zal gaan samenwerken met **Sara**, die ondertussen een goede vriendin is geworden. Het klikt wel tussen ons, dus ik heb volledig het vertrouwen in onze samenwerking. Ook al zijn we allebei leeuwtjes - het zal er dus af en toe tegen zitten - we gaan er samen iets van maken! Voeg daarbij nog iemand zoals **Harald**, de wervelwind, en het is af. Harald, ook al ben je een geweldige arts, verlies ons en de research niet uit het oog! Bedankt voor het verzamelen van de honderden stalen en gegevens...
Vandaag heeft **Hilde** me er meermaals op gewezen dat ik toch eens zal moeten verder schrijven. Hilde ga ik zeker missen, ze is een geweldige steun en luisterend oor. Ik bedenkt zonet dat ik misschien niet erg doordoe, omdat dit het afsluiten betekent van een periode waarvan ik echt heb genoten. Er is bijna geen dag geweest dat ik niet graag ging werken. Ik herinner mij Montreal. Samen met **Ken, Jeroen** en **Pieter** zitten we op een terras, en we drinken een pintje. Wat een fijn congres was dat! Ik kan niet goed tegen verandering, dat is altijd zo geweest. Nostalgie...
Er is zoveel gebeurd, zoveel gezegd, en ik heb zoveel geleerd. Ik heb dan ook veel nood aan discussie, net zoveel als dat ik nood heb om aan wetenschap te doen. De discussie is belangrijk, en ik vrees dat ik er meerderen mee heb verveeld. Vandaar dat dit tegelijk een "sorry, maar toch merci" moet worden.
Pieter, als men ooit een een positivo-gen gaat cloneren, zal jij de power zeker en vast verhogen. Bedankt voor de begeleiding, de steun en de immer positieve sfeer. Trouwens, ik wens jou, maar natuurlijk ook **Lothar** en **Sabine**, veel succes in het nieuwe bedrijf, en verzorg ons Hilde en Inge daar als prinseskes. **Inge**, de reddende engel en paniek-demper, altijd in de buurt om te helpen als er iets misgaat, en er is wat misgegaan, he! Weet je nog dat ik mijn cellen op de grond liet vallen? Gewoon opzuigen met een pasteurpipet en verder doen, zei ze. Het was nog goed afgelopen ook...

Erik Remaut zal ik nooit vergeten. Iemand (ik weet natuurlijk niet meer wie) zei ooit: "*Active science narrows the mind more often than it broadens it*", iets dat ik probeer in gedachte te houden. Erik is het levende bewijs dat dit niet altijd hoeft te gelden. Ik heb een mateloos respect voor wetenschappers die niet door allerlei systemen zijn aangetast, en die buiten een groot verstand, ook een groot hart hebben. Erik is ook de man die in 1 zin (al is die soms lang) kritisch kan samenvatten, en het voorbije uur vergaderen overbodig had kunnen maken. Erik, ik wens je al het allerbeste, geniet van je pensioen.

En dan ons **Michaela**, dedju toch, ons Michke. Waarom toch, he? Hoe komt het dat als er één persoon is op de wereld die absoluut een doctoraat moet verdienen, juist die er het hardst voor moet knokken. Je bent dan ook nog eens een mooi mens, verlies de moed niet, het tij zal wel keren... Samen met ons Karolien moeten je nog unit 1 recht houden, he! Veel plezier, en als je nog eens aan mij denkt, ik ben altijd bereid om gezellig samen te gaan eten... **Karolien**, onze gesprekjes over de zo vaak gelijkaardige "ingebeelde" kwaaltjes en angsten, en natuurlijk over de poes zal ik missen. Nog veel succes met je doctoraat, hopelijk gauw jou feestje? Ik denk nu ook aan **Veerle**. Veel geluk met je nieuwe kleine "projectje", en natuurlijk ook met al de rest. Ons Ve, altijd goedlachs en fijn om mee te kletsen. Dat geldt natuurlijk ook voor **An**, altijd *in the mood* om eens goed te lachen, en altijd bereid om te helpen met een probleem. **Klaas** en **Sofie**, nog veel geluk samen, en, Sofie, ik hoop dat je snel een fijne job vindt! Ik vergeet nog bijna onze **Koen**! 'T Is ook zo'n stilleke, alhoewel... Stille waters? Veel succes op je nieuwe job, en ik hoop dat er gauw een tof vrouwke passeert die je kan thuis houden!

Dan zijn er nog de twee Tommen. **Tombo**, ik hoop dat je de moed vindt om eens eindelijk al je data samen te bundelen in een doctoraat, want dat is nu het enige dat telt voor jou. Je hebt zoveel in je mars. Ik heb veel aan jou te danken, zeker in het begin van mijn doctoraat. Denk nu eerst aan jezelf. **Tom** is ondertussen vertrokken naar het buitenland. Ik hoop dat alles daar verloopt zoals gepland en dat hij daar met zijn vrouwtje een fantastische tijd beleeft. Tom, bedankt voor de leuke tijd! **Steven**, ik ben je inderdaad zeer dankbaar dat je mijn data hebt gered op een kritiek moment. **Miekje**, bedankt voor al die sequenties en hulp. Verder mag ik zeker niet vergeten om de **GOA** te bedanken. Deze maandelijkse vergaderingen waren erg belangrijk in het verloop van mijn project. Ook de **microarray faciliteit**, **applied-maths** en de **genetica faciliteit** hebben heel wat van het werk lichter gemaakt. Ondertussen beginnen we met *the girls* van de **pathologie** wat nauwer samen te werken. Ik denk dat we elkaar nog veel gaan nodig hebben, en **Anouk**: de metallothioneines pakken we samen aan he!

De laatste jaren heb ik vaak de opmerking gekregen dat ik wat meer moet buiten komen. **Liesbetje**, bedankt voor het nalezen van mijn tekstjes, en ik hoop dat we binnenkort elkaar terug

wat meer zien! Dat geldt ook voor mijn lieve vriendin **Babske**. Die sauna-dagen moeten we meer doen, of... onze mannen mogen meer verrassingen organiseren!
Chrisje, ik heb zo de indruk dat je hebt geleerd om te gaan met mijn verschrikkelijke gemoedsschommelingen ten tijde van stress. Bedankt daarvoor, en voor de steun en knuffels!
Terwijl ik mij gedurende de laatste vijf jaar heb afvraagd wanneer ik aan kinderen moet beginnen, heeft mijn lieve broer en zijn **Anneke** er twee op de wereld gezet. Eerst “miniatuur Dirk” **Milan** en twee jaar later “nieuwe man” **Patje**. Wat een plezier die twee brengen is geweldig. Mercikes, Dirk en Anneke, dat jullie er zijn.
Mijn lieve **bomma**, die zo trots is op mij, moet ik ook beloven om eens wat meer langs te komen. Ik ga daar zeker werk van maken!

En dan, he... **papa** en **mama**... Naast het feit dat jullie een fantastische dochter op de wereld hebben gezet ☺, is er zoveel om jullie voor te bedanken. Ik denk nog altijd dat ik de beste ouders heb die er ooit hebben bestaan. Papa, de cliche dat “mijne papa alles kan” is zo waar. Ik kan altijd bij jou terecht, ik kijk echt op naar jou. Mama, ik weet dat je erg fier bent op mij. Bedankt voor de steun, liefde en inspiratie...

Juni 2006