Genetic studies in rheumatoid arthritis

Kurreeman, F.

Genetic Studies in Rheumatoid Arthritis

Fina Kurreeman

Genetic Studies in Rheumatoid Arthritis

PROEFSCHRIFT

ter verkrijging van

de graad Doctor aan de Universiteit Leiden,

op gezag van de Rector Magnificus Prof.mr.P.F.van der Heijden,

volgens besluit van het College voor Promoties

te verdedigen op woensdag 4 November 2009

klokke 13.45 uur

door

Fina Kurreeman

Geboren te Plaines Wilhems, Mauritius

in 1979

PROMOTIECOMMISSIE

Promotor:	Prof. Dr. T.W.J.Huizinga
Co-promotor:	Dr. R.E.M. Toes
Overige leden:	Dr. Robert Plenge (Harvard Medical School, Boston,USA)
	Dr. B.P.C.Koeleman(University Medical Center Utrecht)
	Prof.Dr. W. Jukema
	Prof.Dr. M.R.Daha
	Prof.Dr. F.Koning

ISBN: 978-90-9024653-6

Printed by Propress

The printing of this book was financially supported by:

Reumafonds
Sterk in beweging

Contents

Chapter 1

General Introduction

DeoxyriboNucleic Acid, DNA, a substance of high molecular weight, was identified in 1871 by a young Swiss scientist, Friedrich Miescher[1]. Decades later, in one of the culminating points in biology of all time, James Watson and Francis Crick cracked the DNA code[2]. The two strands of the double helix are anti-parallel, with a sugar phosphate backbone on the outside while bases on the inside form the rung of the ladder. Each rung is composed of two base pairs. Either an adenine-thymine (A-T) pair that form a two-hydrogen bond together, or a cytosine-guanine (C-G) pair that form a three-hydrogen bond.

Nearly forty years later, a new defining moment was attained when the Human Genome Project was initiated in 1990. It was successfully completed in 2003, a year which marks the 50th anniversary of the discovery of the double helix structure of DNA by Nobel Prize winners, James Watson and Francis Crick. The overall result was the generation of a high-quality reference DNA sequence for the human genome‘s 3.2 billion base pairs.

Available to researchers worldwide, the human genome reference sequence provides a magnificent and unprecedented biological resource that serves as a basis for research and discovery and, ultimately, a number of practical applications. In 2002, it inspired a consortium of researchers to embark on the HapMap project to characterize all single nucleotide polymorphisms (SNPs) which are single base differences between individuals. By the end of 2005, researchers had created a map of these patterns across the genome by determining the genotypes of one million or more sequence variants, their frequencies and the degree of association between them, in DNA samples from populations with ancestry from parts of Africa, Asia and Europe[3]. In 2007, this consortium reported the completion of over 3 million SNPs, which represents a third of the estimated 10 million SNPs in the human genome[4]. Together, these research milestones have successfully formed the foundation of current genetic and genomic concepts.

The central goal of human genetics is to understand the inherited basis of human variation, not only in determining differences in phenotypes between individuals but also in elucidating predisposition to disease. With these tools in hand, we can now discover sequence variants that affect common disease, facilitate the development of diagnostic tools and enhance our ability to choose targets for therapeutic intervention.

This thesis focuses on investigating genetic risk factors in rheumatoid arthritis (RA), an autoimmune disease with unknown etiology. While there is a clear genetic component to the development of RA[5], unraveling the genes predisposing to this complex disease, as well as other autoimmune diseases, has been rather elusive. In this chapter, I will provide an overview on rheumatoid arthritis, its genetic basis and the genes identified prior to the start of our research.

Rheumatoid Arthritis

Rheumatoid Arthritis (RA) (MIM 180300 [OMIM]) is a chronic autoimmune disease that affects approximately 0.5-1% of the adult population worldwide and is associated with significant disability and early mortality[5]. Patients suffer from inflammation of the synovial membrane that covers the joint. Joints become red, swollen and tender, and stiffness prevents their use. By definition, RA affects multiple joints. Most commonly, small joints of the hands and feet are affected, although larger joints like the shoulder and knee can also be involved, differing per individual. Eventually, synovitis leads to erosion of the joint surface, causing deformity and loss of function[6]. While the causes of early mortality in RA are not well-known, it may be explained by several factors, e.g chronic exposure to inflammation and cumulative toxicity of immunosuppressive drugs resulting in an increased risk to infectious, cardiovascular, gastrointestinal and respiratory disease[7].

Besides these deleterious consequences for the individual patient, there is a considerable socio-economic impact leading to direct and indirect costs of many billion Euros per year. The total cost of the disease in 2006 was estimated at €45 billion in Europe and €42 billion in the United States[8]. It is therefore of the utmost importance to understand the basis of this disease so as to prevent exorbitant socio-economic costs, but primarily to alleviate patient suffering.

Diagnosis and Clinical phenotypes of RA

The diagnosis of RA is based upon a set of clinical and laboratory measures as defined by the American College of Rheumatology (Table 1). Once a diagnosis is established, the disease course of RA patients remains highly variable, ranging from mild symptoms to chronic inflammation and extensive joint damage.

Criterion	Definition
1. Morning stiffness	Morning stiffness in and around the joints, lasting at least 1 hour before maximal improvement
2. Arthritis of 3 or more joint areas	At least 3 joint areas simultaneously have had soft tissue swelling or fluid (not bony overgrowth alone) observed by a physician. The 14 possible areas are right or left PIP, MCP, wrist, elbow, knee, ankle, and MTP joints
3. Arthritis of hand joints	At least 1 area swollen (as defined above) in a wrist, MCP, or PIP joint
4. Symmetric arthritis	Simultaneous involvement of the same joint areas (as defined in 2) on both sides of the body (bilateral involvement of PIPs, MCPs, or MTPs is acceptable without absolute symmetry)
5. Rheumatoid nodules	Subcutaneous nodules, over bony prominences, or extensor surfaces, or in juxtaarticular regions, observed by a physician
6. Serum rheumatoid factor	Demonstration of abnormal amounts of serum rheumatoid factor by any method for which the result has been positive in <5% of normal control subjects
7. Radiographic changes	Radiographic changes typical of rheumatoid arthritis on posteroanterior hand and wrist radiographs, which must include erosions or unequivocal bony decalcification localized in or most marked adjacent to the involved joints (osteoarthritis changes alone do not qualify)

Table1. American College of Rheumatology (ACR) 1987 revised criteria for the classification of Rheumatoid Arthritis. A patient is diagnosed with rheumatoid arthritis if he/she has satisfied at least 4 of the 7 criteria described below. Criteria 1 through 4 must have been present for at least 6 weeks. Patients with 2 clinical diagnoses are not excluded.

RA patients who fulfill the ACR criteria can be divided into two main subsets; those who possess circulating autoantibodies “autoantibody positive” and those who do not “autoantibody negative”. In the context of this thesis, the term autoantibody positive or negative will refer to the presence or absence of either of the two autoantibodies that play a major role in RA, namely, Rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPA). The classical autoantibody associated with RA is RF, an autoantibody directed against the Fc part of immunoglobulin G. RF is not unique to RA, and is present in other autoimmune diseases, infectious diseases and healthy (elderly) individuals. RF is found in 60-70% of RA patients[9]. In contrast, ACPAs are antibodies directed against citrullinated proteins. Citrullination is the post-translational modification of protein-bound arginine into the non-standard amino acid citrulline. This process results in a small change in molecular mass and the loss of a positive charge[10]. These autoantibodies appear early in RA and can be detected years before disease onset[11,12]. While being found in 50–70% of patients, ACPAs display a unique specificity for RA and are rarely detected in other diseases or in healthy controls[10,12,13]. The relevance of these autoantibodies in disease is exemplified by the fact that patients harboring these autoantibodies generally have a more severe disease course[14,15]. However, whether either of these two autoantibodies are part of a mechanism of disease initiation is still unclear. In contrast, no doubt exists that these autoantibodies represent a very useful tool in both diagnostic and prognostic terms[16], as well as in defining a more homogeneous subset of patients to enhance the discovery of risk factors involved[17].

The role of environment and genetics in RA

RA is considered to be a complex disease, and although the full etiology remains unclear, it is widely accepted that interrelated contributions from environmental and genetic factors play a major role[18]. Interestingly, environmental risk factors so far encompass age, gender, smoking, pregnancy, infections, diet and weather[19] (reviewed in Kobayashi et al). While there is an elevated incidence with an increase in age[20], the female to male ratio of RA patients is ~3:1[5], and generally smokers have a higher propensity to develop RA[21].

There is however a rather large genetic component to RA. Evidence from twin studies demonstrates excess disease concordance between monozygotic (15%) when compared with dizygotic (3.6%) twins(22). From such studies, the genetic contribution to RA has been estimated between 50% and 60%[23]. The increased risk of disease in siblings of patients with RA compared with that of the general population (λs) has been estimated to be between 2 and 17 fold[24]. These data altogether provide compelling evidence of the role of genetics in the development of RA.

HLA, the most prominent genetic risk factor in RA

The most prominent genetic association is confined to the human leukocyte antigen (HLA) locus on the short arm of chromosome 6. HLA-DR gene variants have been consistently associated with RA across several populations and in microsatellite-based whole genome screens on affected sibling pair families in Europe, US, UK and Japan[25-31]. This method involves a search for increased sharing of particular genetic regions among affected siblings. The association of HLA with RA has been mapped to the third hypervariable region of DRβ-chains, especially aa 70–74, encoding a conserved amino acid motif (QKRAA, QRRAA, or RRRAA). This susceptibility epitope, called the shared epitope (SE), is found in multiple RA-associated DR molecules, including *DR1, DR4*, *DR10* and *DR14* (i.e *DRB*0101, DRB*0102, DRB*0401*, *DRB*0404*, *DRB*0405, DRB*0408, DRB*1001 and DRB*1402*)(32). However, amino acids encoding the DERAA motif (i.e. DRB*0103, DRB *0402, DRB *1102, DRB *1103, DRB *1301, DRB *1302 and DRB *1304) at the same position have a protective effect on the development of RA[33-36].

The HLA region is gene-rich consisting of over 100 immune-related genes that could be potentially relevant to the pathogenesis of RA[37-39]. Therefore, understanding of the biological role of this region in RA remains to be discovered. However, since the association with the HLA region only accounts for approximately 30% of the genetic burden to RA, it implies that additional genetic risk factors play a role in RA[40].

Non-HLA genes in RA

A number of markers outside the HLA region did emerge from these four microsatellite-based whole genome studies which were suggestive of linkage with RA, although the effect of the HLA region is by far the strongest. One problem is that such studies have weak power to detect modest effects. In contrast, because such studies are simultaneously testing a large number of regions which may be linked with disease, the likelihood of a false-positive result is very high. Regions of linkage also tend to encompass large genetic regions containing a large number of genes making it difficult for researchers to progress from the linkage region to the causal gene. It is therefore not surprising that studies often failed to replicate their results.

One of the few success stories came from a Japanese group. Following indications of linkage from the overlapping locus on chromosome 1p36 among certain genome-wide scans, Yamamoto and colleagues observed that this genomic region contains a cluster of enzymes that is functionally associated with the production of rheumatoid arthritis–specific autoantibodies. These enzymes are the peptidylarginine deiminases (PADIs), which posttranslationally convert arginine residues to citrulline. Fine-mapping of this region containing four of these enzymes (PADI 1-4) located next to one another revealed that polymorphisms in the *PADI4* gene are strongly associated with RA in the Japanese[41]. However, whether this gene is associated with RA in Caucasians remains a question of debate and so far, no conclusive evidence has been obtained[42,43].

Another major breakthrough, employing a large-scale approach did not come from classical linkage studies but association studies between patients and healthy individuals in 2004. Begovich and colleagues performed a large-scale screen utilizing putative functional SNPs and identified a non-synonymous SNP (R620W) in the gene, protein tyrosine phosphatase non-

receptor type 22 (*PTPN22*) more frequently in patients than in healthy individuals[44]. Well powered studies have successfully replicated the same association of rheumatoid arthritis and the R620W polymorphism, in populations of European descent from the UK, Finland, Sweden, Germany, Netherlands, Spain and Canada[45] (reviewed by Bowes *et al*, 2008). Such consensus was previously unprecedented. Intriguingly, both HLA-SE alleles and *PTPN22* are associated with the development of ACPA positive disease(17). It is of note that the PTPN22 R620W SNP is not polymorphic in the Japanese population and a haplotype analysis of the region reveals no association[46].

The same *PTPN22* polymorphism has also been associated with several autoimmune diseases in Caucasians including among others Juvenile Idiopathic Arthritis and Systemic Lupus Erythematosus, Graves disease and Addison's disease[47-49], indicating the genetic risk factors may not be unique to a specific disease but can be promiscuously associated with other autoimmune diseases.

An alternative approach to genome-wide strategies is to use a candidate gene screen which takes a hypothesis-driven approach. While this strategy has generated a huge amount of literature, it has not been very fruitful in the identification of consistent and replicable risk factors for RA outside the HLA region. With the advent of new technologies in genotyping a large number of variants and the availability of SNP data from the HapMap consortium, considerable progress has been made in this field, enabling researchers to perform highly improved association studies.

Outline of this thesis

The identification of RA-associated genes outside of the HLA region has been a challenge. Although the expected effect of genetic factors outside the HLA region are modest, the identification of risk loci through human genetic studies offers *prima facie* evidence that a biological pathway is critical in disease pathogenesis. Therefore, the aim of this thesis was to take a candidate gene approach to identify risk factors involved in rheumatoid arthritis. It is divided into three parts in which **part one** is dedicated towards the investigation of a region of the genome encompassing genes highly involved in the immune system, namely *Tumour Necrosis Factor (TNF) Receptor associated factor 1/Complement component 5 (TRAF1/C5)* on chromosome 9q33. In the **second part**, we have investigated the role of an immunoregulatory cytokine interleukin 10 *(IL10)* located on chromosome 1q32 and in **part three** we have investigated the role of additional genetic risk factors in RA.

In **the first part** of the thesis, we have investigated the role of the *TRAF1/C5* region in RA as well as other autoimmune diseases. **In chapter 2**, we have described the *TRAF1/C5* region as one of the few widely-replicated genetic risk factors for RA. Based on available information in mouse models[50-52] and indications from human studies[53], we hypothesized that the *C5* region may play a role in the development and/or exacerbation of arthritis. By genetic fine-mapping studies, we identified the haplotype associating with disease and replicated our findings in 4 different cohorts derived from three different populations from the Netherlands, US and Sweden. Intriguingly, a genome wide association study (GWAS) on SNPs performed by Plenge *et al* identified the same *TRAF1/C5* region, in addition to the previously known *HLA* and *PTPN22*, as genetic risk factors for RA[54].

To further establish this risk factor, we have reproduced this association in trio families in which both parents are unaffected and one offspring affected with RA **(Chapter 3)**. This locus represents the third genetic risk factor for which association is found in family-based studies as well. Together, these findings firmly establish the *TRAF1/C5* region as the one of the confirmed genetic risk factors for RA.

In the recent years, it has been suggested that there may be a considerable heritable component to autoimmune diseases[55]. While certain diseases such as RA tend to occur among several members of the same family indicating a genetic component to that specific disease, it is also common to observe different autoimmune diseases in various family members as well as in a particular individual, suggesting that certain individuals may have inherited a set of genetic risk factors predisposing them to the development of an autoimmune disease. Our work has now shown that the *TRAF1/C5* region is not only relevant for RA but is also relevant in patients with a polyarticular form of juvenile arthritis (JIA) (**Chapter 4**). Behrens and colleagues have now also independently reported an association of a perfect proxy to JIA further supporting our findings[56].

To test whether this region predisposes to other diseases we also investigated its relevance in a well-powered study including four additional autoimmune diseases including Type I Diabetes (TID), Celiac Disease (CD), Systemic Sclerosis (SSc), Systemic Lupus Erythematosus (SLE) patients and a common set of controls consisting of healthy unrelated individuals that were geographically and ethnically matched. We observe and replicate modest associations to both T1D and SLE and did not observe any evidence of association to CD and SSc (**Chapter 5**).

With these studies, we have provided considerable evidence that the *TRAF1/C5* region is not only relevant to RA but that the frequency of the same allele is increased in JIA, T1D and SLE. It is therefore highly likely that the *TRAF1/C5* region is a genetic risk factor involved in a shared pathway underlying multiple autoimmune diseases.

The **second part** of this thesis addresses the role of *interleukin 10* (*IL10*) genetic variants in regulating expression levels and their role in disease. IL10 is a cytokine with key regulatory, anti-inflammatory and immuno-stimulatory functions involved in the pathogenesis of various diseases[57-60]. Interindividual differences in the production of IL10 have been extensively associated with polymorphisms and haplotypes of the *IL10* gene[61]. The A allele of IL10-2849, a polymorphism located in the promoter region, is associated with decreased IL10 production as measured by lipopolysaccharide (LPS) stimulated whole blood cultures[62]. A low innate production of IL10 using the same assay is associated with an increased risk of familial osteoarthritis (OA)[63]. Therefore, in **chapter 6**, we investigate the role of *IL10* in osteoarthritis and observe no association of 7 promoter SNPs with disease.

While there is a direct correlation between IL10 mRNA and protein levels and high and low IL10 producers have similar mRNA halflife(64), little evidence existed that this variation in clinically relevant levels of IL10 is actually dictated by *IL10* haplotypes. In **chapter 7**, by using the technique of allele-specific transcript quantification (ASTQ), the ratio between two alleles (A and G) of the *IL10* gene was characterized in 15 healthy heterozygous individuals. We show that *IL10* alleles are indeed differentially transcribed in cells from heterozygous individuals and that *IL10* haplotypes likely dictate the production of IL10. These findings show, for the first time, that interindividual differences in IL10 protein levels could be partially explained at the allele-specific transcriptional level.

In RA, the IL10-A2849 G allele has been shown to be associated with differences in titres of autoantibodies (RF and ACPA). Moreover the rate of joint destruction in RA patients from the early arthritis cohort was twice as high in patients that were –2849G carrier to those who were not (median rate per year 8 versus 4 SHS units on X-rays of hand and feet)(65). As the length of the haplotype block around *IL10* is highly relevant to the search for the functional polymorphism(s), we characterized the level of linkage disequilibrium in a region of 217 kb, encompassing *IL10* as well as its neighbouring homologues (*IL19, IL20* and *IL24*). We showed that the neighboring genes are unlikely to harbor functional cis-acting variants (**chapter 8**). In this chapter we also report an association of IL10 polymorphisms with restenosis. Stenosis occurs when a coronary artery constricts or narrows. One way to widen a coronary artery is by using percutaneous coronary intervention (PCI, or balloon angioplasty). Some patients who undergo PCI have restenosis (renarrowing) of the widened segment within about six months of the procedure. Restenosed arteries may have to undergo another angioplasty. Inflammation is thought to play a key role in the development of restenosis and concordantly, we observed that patients with a lower innate ability to make IL10, favouring a pro-inflammatory environment, are at a high risk of undergoing restenosis.

We further fine-mapped the immediate *IL10* region. Six tagging SNPs have been genotyped in our extensive and clinically well-defined RA cohorts to determine their relevance to clinical remission and severity of RA (**Chapter 9**). While there is conflicting evidence that the *IL10* gene is associated with the development of RA, there is no indication in our cohort of the involvement of *IL10* in either clinical remission or the progression of joint destruction in RA.

In the **third part** of this thesis, the role of other genetic risk factors (besides *IL10* and *TRAF1/C5*) is described. In **chapter 10**, the role of *TNFα* in predisposing patients towards a more severe disease course is investigated. While increased levels of TNFα are found in patients with RA[18] and the treatment of patients with anti-TNF agents do provide beneficial effects[66], very little evidence exists that variations in the gene predispose individuals to the development or progression of rheumatoid arthritis, implying that increased TNF production in RA patients is most likely due to other molecules in the signaling cascade leading to the enhanced production of TNFα protein. One such molecule *TNFα-induced protein 3* (*TNFAIP3*) on chromosome 6q23 has recently been associated with development of RA(67;68). *TNFAIP3* is a negative regulator of *NFκB* and as such is involved in inhibiting TNF-Receptor mediated signaling effects[69]. Interestingly, the initial association was detected in cohorts of patients with long-standing RA. However, no association was found in a Swedish early arthritis cohort. We therefore hypothesized that the 6q23 locus containing *TNFAIP3* may be predominantly associated with a phenotype consistent with more severe disease. To determine whether this is the case, we set out in **chapter 11** to analyze the effect of the 6q23 region on the rate of joint destruction in our large and well-described early RA cohort.

One of the rare success stories for RA from the classical microsatellite-based linkage approach came in 2007, when candidate genes were investigated under a linkage peak on chromosome 2q in the US study. The 13 candidate genes investigated revealed strong association at four strongly linked SNPs in an intron of the gene *signal transducer and activator of transcription 4* (*STAT4)*[67]. The data from *STAT4* (Chr2q33) has already been consistently replicated in not only Caucasians but also in East Asians[68-71]. This is however not the case for the other signal observed under the same linkage peak in this study, namely cytotoxic T lymphocyte associated 4 gene (*CTLA4*). Originally identified as a determinant of susceptibility to autoimmune diseases including Grave's disease, Type 1 Diabetes and autoimmune hypothyroidism[72], this locus has been a constant debate in RA[42]. In **chapter 12**, we perform independent replication and a meta-analysis of three loci including *STAT4*, *CTLA4* and the recently described 4q27 region containing the *IL2* and *IL21* genes[73]. We show a strong association with *STAT4* and as previously described, no preferential association was observed with ACPA status. More importantly, we confirm the role of *CTLA4* in RA, resolving a longstanding debate of whether does or does not predispose to RA. We additionally show for the first time that the association is restricted to ACPA positive individuals only. For the 4q27 locus, we provide independent replication of the data and indicate that for this locus, in contrast to previous findings, no differences in effects are seen in ACPA positive and ACPA negative individuals. However, these data have to be interpreted with caution due to a possible lack of power. Interestingly, we also observe an association between the 4q27 locus and juvenile arthritis as described in **chapter 13**.

While most described genetic risk factors in RA either predispose to the autoantibody positive subset of patients or both, data is extremely scarce when it comes to patients who harbor none. One genetic risk factor in the HLA region, DRB*0301, has been consistently associated with ACPA negative disease[74,75]. In **chapter 14**, we describe the identification of the only non-HLA genetic factor, *Interferon Regulatory factor 5 (IRF5),* showing a predominant association with ACPA negative disease. A recent report confirmed these findings but also show a small effect in ACPA positive disease[76].

Chapter 15 provides an overview of novel genetic risk factors in ACPA positive RA, identified with the use of a meta-analysis of three well-powered GWAS studies from the US, Sweden and the UK. Simultaneously, the UK group confirmed 3 of the loci identified in the meta-analysis including MMEL1, PCRKQ and KIF5A[80]. Two additional loci surfaced in their study providing compelling evidence for IL2RB and suggestive evidence for IL2RA. In **Chapter 16**, we provide the first independent study replicating these two risk factors in a non-UK population, underlining the relevance of the IL2 pathway in RA. Finally in **Chapter 17**, I discuss the biological relevance and potential implications of all these identified RA loci, summarizing the recent explosion of genetic findings in RA.

References

(1) Pollock MR. The discovery of DNA: An ironic tale of chance, prejudice and insight. J Gen Microbiol 1970; 63(1):1-20.

(2) WATSON JD, CRICK FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 1953; 171(4356):737-8.

(3) A haplotype map of the human genome. Nature 2005; 437(7063):1299-320.

(4) Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007; 449(7164):851-61.

(5) Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. Arthritis Res 2002; 4 Suppl 3:S265-S272.

(6) Majithia V, Geraci SA. Rheumatoid arthritis: diagnosis and management. Am J Med 2007; 120(11):936-9.

(7) Gabriel SE, Crowson CS, Kremers HM, Doran MF, Turesson C, O'Fallon WM et al. Survival in rheumatoid arthritis: a population-based analysis of trends over 40 years. Arthritis Rheum 2003; 48(1):54-8.

(8) Lundkvist J, Kastang F, Kobelt G. The burden of rheumatoid arthritis and access to treatment: health burden and costs. Eur J Health Econ 2008; 8 Suppl 2:S49-S60.

(9) Visser H, Gelinck LB, Kampfraath AH, Breedveld FC, Hazes JM. Diagnostic and prognostic characteristics of the enzyme linked immunosorbent rheumatoid factor assays in rheumatoid arthritis. Ann Rheum Dis 1996; 55(3):157-61.

(10) van Gaalen F, Ioan-Facsinay A, Huizinga TW, Toes RE. The devil in the details: the emerging role of anticitrulline autoimmunity in rheumatoid arthritis. J Immunol 2005; 175(9):5575-80.

(11) Nielen MM, van Schaardenburg D, Reesink HW, van de Stadt RJ, van der Horst-Bruinsma IE, de Koning MH et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004; 50(2):380-6.

(12) Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003; 48(10):2741-9.

(13) van Gaalen FA, Visser H, Huizinga TW. A comparison of the diagnostic accuracy and prognostic value of the first and second anti-cyclic citrullinated peptides (CCP1 and CCP2) autoantibody tests for rheumatoid arthritis. Ann Rheum Dis 2005; 64(10):1510-2.

(14) De Rycke L, Peene I, Hoffman IE, Kruithof E, Union A, Meheus L et al. Rheumatoid factor and anticitrullinated protein antibodies in rheumatoid arthritis: diagnostic value, associations with radiological progression rate, and extra-articular manifestations. Ann Rheum Dis 2004; 63(12):1587-93.

(15) van der Helm-van Mil AH, Verpoort KN, Breedveld FC, Toes RE, Huizinga TW. Antibodies to citrullinated proteins and differences in clinical progression of rheumatoid arthritis. Arthritis Res Ther 2005; 7(5):R949-R958.

(16) Bugatti S, Codullo V, Caporali R, Montecucco C. B cells in rheumatoid arthritis. Autoimmun Rev 2007; 7(2):137-42.

(17) van der Helm-van Mil AH, Huizinga TW, de Vries RR, Toes RE. Emerging patterns of risk factor make-up enable subclassification of rheumatoid arthritis. Arthritis Rheum 2007; 56(6):1728-35.

(18) Firestein GS. Evolving concepts of rheumatoid arthritis. Nature 2003; 423(6937):356-61.

(19) Kobayashi S, Momohara S, Kamatani N, Okamoto H. Molecular aspects of rheumatoid arthritis: role of environmental factors. FEBS J 2008.

(20) Goronzy JJ, Weyand CM. Rheumatoid arthritis. Immunol Rev 2005; 204:55-73.

(21) Klareskog L, Padyukov L, Ronnelid J, Alfredsson L. Genes, environment and immunity in the development of rheumatoid arthritis. Curr Opin Immunol 2006; 18(6):650-5.

(22) Silman AJ, MacGregor AJ, Thomson W, Holligan S, Carthy D, Farhan A et al. Twin concordance rates for rheumatoid arthritis: results from a nationwide study. Br J Rheumatol 1993; 32(10):903-7.

(23) MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000; 43(1):30-7.

(24) Seldin MF, Amos CI, Ward R, Gregersen PK. The genetics revolution and the assault on rheumatoid arthritis. Arthritis Rheum 1999; 42(6):1071-9.

(25) Cornelis F, Faure S, Martinez M, Prud'homme JF, Fritz P, Dib C et al. New susceptibility locus for rheumatoid arthritis suggested by a genome-wide linkage study. Proc Natl Acad Sci U S A 1998; 95(18):10746-50.

(26) Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Monteiro J et al. A genomewide screen in multiplex rheumatoid arthritis families suggests genetic overlap with other autoimmune diseases. Am J Hum Genet 2001; 68(4):927-36.

(27) MacKay K, Eyre S, Myerscough A, Milicic A, Barton A, Laval S et al. Whole-genome linkage analysis of rheumatoid arthritis susceptibility loci in 252 affected sibling pairs in the United Kingdom. Arthritis Rheum 2002; 46(3):632-9.

(28) Shiozawa S, Hayashi S, Tsukamoto Y, Goko H, Kawasaki H, Wada T et al. Identification of the gene loci that predispose to rheumatoid arthritis. Int Immunol 1998; 10(12):1891-5.

(29) Eyre S, Barton A, Shephard N, Hinks A, Brintnell W, MacKay K et al. Investigation of susceptibility loci identified in the UK rheumatoid arthritis whole-genome scan in a further series of 217 UK affected sibling pairs. Arthritis Rheum 2004; 50(3):729-35.

(30) Etzel CJ, Chen WV, Shepard N, Jawaheer D, Cornelis F, Seldin MF et al. Genome-wide meta-analysis for rheumatoid arthritis. Hum Genet 2006; 119(6):634-41.

(31) Jawaheer D, Seldin MF, Amos CI, Chen WV, Shigeta R, Etzel C et al. Screening the genome for rheumatoid arthritis susceptibility genes: a replication study and combined analysis of 512 multicase families. Arthritis Rheum 2003; 48(4):906-16.

(32) Zanelli E, Breedveld FC, de Vries RR. HLA class II association with rheumatoid arthritis: facts and interpretations. Hum Immunol 2000; 61(12):1254-61.

(33) van der Helm-van Mil AH, Huizinga TW, Schreuder GM, Breedveld FC, de Vries RR, Toes RE. An independent role of protective HLA class II alleles in rheumatoid arthritis severity and susceptibility. Arthritis Rheum 2005; 52(9):2637-44.

(34) de Vries N, Tijssen H, van Riel PL, van de Putte LB. Reshaping the shared epitope hypothesis: HLA-associated risk for rheumatoid arthritis is encoded by amino acid substitutions at positions 67-74 of the HLA-DRB1 molecule. Arthritis Rheum 2002; 46(4):921-8.

(35) Shadick NA, Heller JE, Weinblatt ME, Maher NE, Cui J, Ginsburg G et al. Opposing effects of the D70 mutation and the shared epitope in HLA-DR4 on disease activity and certain disease phenotypes in rheumatoid arthritis. Ann Rheum Dis 2007; 66(11):1497-502.

(36) Mattey DL, Dawes PT, Gonzalez-Gay MA, Garcia-Porrua C, Thomson W, Hajeer AH et al. HLA-DRB1 alleles encoding an aspartic acid at position 70 protect against development of rheumatoid arthritis. J Rheumatol 2001; 28(2):232-9.

(37) Horton R, Wilming L, Rand V, Lovering RC, Bruford EA, Khodiyar VK et al. Gene map of the extended human MHC. Nat Rev Genet 2004; 5(12):889-99.

(38) Horton R, Gibson R, Coggill P, Miretti M, Allcock RJ, Almeida J et al. Variation analysis and gene annotation of eight MHC haplotypes: the MHC Haplotype Project. Immunogenetics 2008; 60(1):1-18.

(39) Stewart CA, Horton R, Allcock RJ, Ashurst JL, Atrazhev AM, Coggill P et al. Complete MHC haplotype sequencing for common disease gene mapping. Genome Res 2004; 14(6):1176-87.

(40) Deighton CM, Walker DJ, Griffiths ID, Roberts DF. The contribution of HLA to rheumatoid arthritis. Clin Genet 1989; 36(3):178-82.

(41) Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M et al. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 2003; 34(4):395-402.

(42) Oliver JE, Worthington J, Silman AJ. Genetic epidemiology of rheumatoid arthritis. Curr Opin Rheumatol 2006; 18(2):141-6.

(43) Plenge RM, Padyukov L, Remmers EF, Purcell S, Lee AT, Karlson EW et al. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. Am J Hum Genet 2005; 77(6):1044-60.

(44) Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004; 75(2):330-7.

(45) Bowes J, Barton A. Recent advances in the genetics of RA susceptibility. Rheumatology (Oxford) 2008; 47(4):399-402.

(46) Ikari K, Momohara S, Inoue E, Tomatsu T, Hara M, Yamanaka H et al. Haplotype analysis revealed no association between the PTPN22 gene and RA in a Japanese population. Rheumatology (Oxford) 2006; 45(11):1345-8.

(47) Gregersen PK. Gaining insight into PTPN22 and autoimmunity. Nat Genet 2005; 37(12):1300-2.

(48) Gregersen PK, Lee HS, Batliwalla F, Begovich AB. PTPN22: setting thresholds for autoimmunity. Semin Immunol 2006; 18(4):214-23.

(49) Lee YH, Rho YH, Choi SJ, Ji JD, Song GG, Nath SK et al. The PTPN22 C1858T functional polymorphism and autoimmune diseases--a meta-analysis. Rheumatology (Oxford) 2007; 46(1):49-56.

(50) Firestein GS. Evolving concepts of rheumatoid arthritis. Nature 2003; 423(6937):356-61.

(51) Ji H, Gauguier D, Ohmura K, Gonzalez A, Duchatelle V, Danoy P et al. Genetic influences on the end-stage effector phase of arthritis. J Exp Med 2001; 194(3):321-30.

(52) Ji H, Ohmura K, Mahmood U, Lee DM, Hofhuis FMA, Boackle SA et al. Arthritis Critically Dependent on Innate Immune System Players. Immunity 2002; 16(2):157-68.

(53) Hogasen K, Mollnes TE, Harboe M, Gotze O, Hammer HB, Oppermann M. Terminal complement pathway activation and low lysis inhibitors in rheumatoid arthritis synovial fluid. J Rheumatol 1995; 22(1):24-8.

(54) Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B et al. TRAF1-C5 as a risk locus for rheumatoid arthritis - A genomewide study. New England Journal of Medicine 2007; 357(12):1199-209.

(55) Gregersen PK, Behrens TW. Genetics of autoimmune diseases--disorders of immune homeostasis. Nat Rev Genet 2006; 7(12):917-28.

(56) Behrens EM, Finkel TH, Bradfield JP, Kim CE, Linton L, Casalunovo T et al. Association of the TRAF1-C5 locus on chromosome 9 with juvenile idiopathic arthritis. Arthritis Rheum 2008; 58(7):2206-7.

(57) de Jong BA, Westendorp RG, Eskdale J, Uitdehaag BM, Huizinga TW. Frequency of functional interleukin-10 promoter polymorphism is different between relapse-onset and primary progressive multiple sclerosis. Hum Immunol 2002; 63(4):281-5.

(58) Lin MT, Storer B, Martin PJ, Tseng LH, Gooley T, Chen PJ et al. Relation of an interleukin-10 promoter polymorphism to graft-versus-host disease and survival after hematopoietic-cell transplantation. N Engl J Med 2003; 349(23):2201-10.

(59) O'Garra A, Vieira P. T(H)1 cells control themselves by producing interleukin-10. Nat Rev Immunol 2007; 7(6):425-8.

(60) Eskdale J, McNicholl J, Wordsworth P, Jonas B, Huizinga T, Field M et al. Interleukin-10 microsatellite polymorphisms and IL-10 locus alleles in rheumatoid arthritis susceptibility. Lancet 1998; 352(9136):1282-3.

(61) Eskdale J, Keijsers V, Huizinga T, Gallagher G. Microsatellite alleles and single nucleotide polymorphisms (SNP) combine to form four major haplotype families at the human interleukin-10 (IL-10) locus. Genes Immun 1999; 1(2):151-5.

(62) Westendorp RG, Langermans JA, Huizinga TW, Verweij CL, Sturk A. Genetic influence on cytokine production in meningococcal disease. Lancet 1997; 349(9069):1912-3.

(63) Riyazi N, Slagboom E, de Craen AJ, Meulenbelt I, Houwing-Duistermaat JJ, Kroon HM et al. Association of the risk of osteoarthritis with high innate production of interleukin-1beta and low innate production of interleukin-10 ex vivo, upon lipopolysaccharide stimulation. Arthritis Rheum 2005; 52(5):1443-50.

(64) Eskdale J, Gallagher G, Verweij CL, Keijsers V, Westendorp RG, Huizinga TW. Interleukin 10 secretion in relation to human IL-10 locus haplotypes. Proc Natl Acad Sci U S A 1998; 95(16):9465-70.

(65) Lard LR, van Gaalen FA, Schonkeren JJ, Pieterman EJ, Stoeken G, Vos K et al. Association of the -2849 interleukin-10 promoter polymorphism with autoantibody production and joint destruction in rheumatoid arthritis. Arthritis Rheum 2003; 48(7):1841-8.

(66) Maini RN, Elliott MJ, Brennan FM, Feldmann M. Beneficial effects of tumour necrosis factor-alpha (TNF-alpha) blockade in rheumatoid arthritis (RA). Clin Exp Immunol 1995; 101(2):207-12.

(67) Plenge RM, Cotsapas C, Davies L, Price AL, Bakker PIW, Maller J et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nature Genetics 2007; 39(12):1477-82.

(68) Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J et al. Rheumatoid arthritis association at 6q23. Nat Genet 2007; 39(12):1431-3.

(69) Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S et al. De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. Nature 2004; 430(7000):694-9.

(70) Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007; 357(10):977-86.

(71) Barton A, Thomson W, Ke X, Eyre S, Hinks A, Bowes J et al. Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility. Hum Mol Genet 2008; 17(15):2274-9.

(72) Kobayashi S, Ikari K, Kaneko H, Kochi Y, Yamamoto K, Shimane K et al. Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population. Arthritis Rheum 2008; 58(7):1940-6.

(73) Orozco G, Alizadeh BZ, Delgado-Vega AM, Gonzalez-Gay MA, Balsa A, Pascual-Salcedo D et al. Association of STAT4 with rheumatoid arthritis: a replication study in three European populations. Arthritis Rheum 2008; 58(7):1974-80.

(74) Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, Gregersen PK. Association of STAT4 with rheumatoid arthritis in the Korean population. Mol Med 2007; 13(9-10):455-60.

(75) Ueda H, Howson JM, Esposito L, Heward J, Snook H, Chamberlain G et al. Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease. Nature 2003; 423(6939):506-11.

(76) Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, Franke B et al. Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. Am J Hum Genet 2007; 81(6):1284-8.

(77) Verpoort KN, van Gaalen FA, van der Helm-van Mil AH, Schreuder GM, Breedveld FC, Huizinga TW et al. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. Arthritis Rheum 2005; 52(10):3058-62.

(78) Irigoyen P, Lee AT, Wener MH, Li W, Kern M, Batliwalla F et al. Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA-DR3 and the shared epitope alleles. Arthritis Rheum 2005; 52(12):3813-8.

(79) Dieguez-Gonzalez R, Calaza M, Perez-Pampin E, de la Serna AR, Fernandez-Gutierrez B, Castaneda S et al. Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis. Arthritis Rheum 2008; 58(5):1264-74.

(80) Barton A, Thomson W, Ke X, Eyre S, Hinks A, Bowes J et al. Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. Nat Genet 2008; 40(10):1156-9.

Part I

Association of the TRAF1-C5 region on chromosome 9q33 with RA and autoimmunity

Chapter 2

A Candidate Gene Approach Identifies the TRAF1/C5 Region as a Risk Factor for Rheumatoid Arthritis

F.A.S. Kurreeman
L. Padyukov
R.B. Marques
S.J. Schrodi
M. Seddighzadeh
G. Stoeken-Rijsbergen
A.H.M. van der Helm-van Mil
C.F. Allaart
W. Verduyn
J. Houwing-Duistermaat
L. Alfredsson
A.B. Begovich
L. Klareskog
T.W.J. Huizinga
R.E.M. Toes

Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands
Rheumatology Unit, Department of Medicine, Karolinska Institute at Karolinska Hospital, Stockholm, Sweden; Celera, Alameda, California, United States of America
Department of Immunohaematology and Bloodbank, Leiden University Medical Center, Leiden, The Netherlands
Institute of Environmental Medicine, Department of Medical Statistics, Leiden University Medical Centre, Leiden, The Netherlands

PLoS Medicine. 4, e278 (2007)

Abstract

Background

Rheumatoid arthritis (RA) is a chronic autoimmune disorder affecting ~1% of the population. The disease results from the interplay between an individual's genetic background and unknown environmental triggers. Although human leukocyte antigens (HLAs) account for ~30% of the heritable risk, the identities of non-HLA genes explaining the remainder of the genetic component are largely unknown. Based on functional data in mice, we hypothesized that the immune-related genes *complement component 5 (C5)* and/or *TNF receptor-associated factor 1 (TRAF1),* located on Chromosome 9q33-34, would represent relevant candidate genes for RA. We therefore aimed to investigate whether this locus would play a role in RA.

Methods and Findings

We performed a multitiered case-control study using 40 single-nucleotide polymorphisms (SNPs) from the *TRAF1* and *C5* (*TRAF1/C5*) region in a set of 290 RA patients and 254 unaffected participants (controls) of Dutch origin. Stepwise replication of significant SNPs was performed in three independent sample sets from the Netherlands ($n_{cases/controls}$ = 454/270), Sweden ($n_{cases/controls}$ = 1,500/1,000) and US ($n_{cases/controls}$ = 475/475). We observed a significant association ($p < 0.05$) of SNPs located in a haplotype block that encompasses a 65 kb region including the 3′ end of *C5* as well as *TRAF1*. A sliding window analysis revealed an association peak at an intergenic region located ~10 kb from both *C5* and *TRAF1*. This peak, defined by SNP14/rs10818488, was confirmed in a total of 2,719 RA patients and 1,999 controls (odds ratio$_{common}$ = 1.28, 95% confidence interval 1.17–1.39, $p_{combined} = 1.40 \times 10^{-8}$) with a population attributable risk of 6.1%. The A (minor susceptibility) allele of this SNP also significantly correlates with increased disease progression as determined by radiographic damage over time in RA patients ($p = 0.008$).

Conclusions

Using a candidate-gene approach we have identified a novel genetic risk factor for RA. Our findings indicate that a polymorphism in the *TRAF1/C5* region increases the susceptibility to and severity of RA, possibly by influencing the structure, function, and/or expression levels of *TRAF1* and/or *C5*.

Introduction

Rheumatoid arthritis (RA) is characterized by chronic inflammation and destruction of the synovial joints leading to progressive joint damage and disability. The disease has a complex etiology, including a wide spectrum of clinical manifestations, variability in disease severity and/or progression, and differential response to a range of therapies. This heterogeneous phenotype suggests the involvement of both environmental and genetic factors [1], where the genetic component of RA has been estimated to be between 50%–60% [2,3]. Identification of disease-associated genes is important as it will guide our understanding of the biological pathways underlying polygenic diseases and the development of potential novel therapeutic targets.

The most prominent genetic association in RA is confined to the *human leukocyte antigen (HLA)* locus. Although this association has been known for almost 30 years, and although the underlying mechanism is still not understood, it has been replicated in multiple studies [2,4]. The identification of RA-associated genes outside of the *HLA* region, however, has been a challenge. Recently one such gene, *protein tyrosine phosphatase, non-receptor type 22 (lymphoid) (PTPN22)*, was identified in the first step of a large genetic-association study utilizing putative functional SNPs [5]. The gene product encoded by *PTPN22* is, like the *HLA* locus, involved in T cell-mediated immune responses. However, other immune components are also thought to play a pivotal role in RA, as demonstrated by the beneficial effects of treatment with agents that block proinflammatory cytokines, such as tumor necrosis factor α (TNFα) [6]. Moreover, in several experimental animal models for RA, innate immune responses mediated by a diversity of players have been implicated in arthritis. In this respect, a prominent role for the complement system has been identified as mice deficient in complement factors are resistant to arthritis, and as it has been shown that targeting *complement component 5 (C5)* by antibodies prevents the onset of arthritis and reduces the clinical severity in mouse models for arthritis [7,8]. Likewise, the observation that high levels of C5a, a potent chemoattractant, are found in synovial fluid of RA patients combined with the fact that *C5a receptor*-deficient mice are also resistant to arthritis induction, indicate a central role for these mediators in arthritis [9,10]. A genome scan of mice that were or were not susceptible to antibody-induced arthritis revealed that the main genetic influence detected in this model maps to the *C5* region [11].

These functional data in mice inspired us to hypothesize that the *C5* region would be a contributing factor in RA. Therefore, we searched for further evidence by first addressing the question of whether any genetic indications exist that implicate the involvement of this region in RA. A conventional linkage study using microsatellite markers identified a linkage peak in the vicinity of the *C5* region [12]. Although this study provided weak evidence for linkage (logarithm of the odds score, LOD 1.8), it further boosted our interest in this region. *C5* is located next to *TNF receptor-associated factor 1 (TRAF1)*, an essential effector of the TNF signaling cascade. Since TNF blockade represents a powerful intervention in both mice and humans for the treatment of arthritis, it provided an additional rationale to explore this genetic region encoding *C5* and *TRAF1,* which are adjacent to each other on Chromosome 9q33-34. We therefore sought to investigate whether these candidate genes, which are important immune mediators, would play a role in RA.

Methods

Study Populations

All RA patients in all sets in this study met the American College of Rheumatology 1987 revised criteria for RA [13]. Sample set 1 cases consisted of 290 RA patients consecutively included from the out-patient clinic of the Leiden University Hospital in 1994 [14] and 254 controls randomly selected by the section Immunogenetics and Transplantation Immunology of Leiden University Medical Center, Leiden, The Netherlands (ITI).

Since 89% of the first set of RA patients were rheumatoid factor (RF) positive, we genotyped an independent sample set 2, which consisted of 454 RF-positive patients from two inception cohorts of early arthritis patients (EAC and BEST) and a second set of 270 randomly selected Dutch blood donors from ITI. Briefly, the EAC consists of patients included from 1993 onwards and originating from a health care region of about 400,000 inhabitants in the western part of The Netherlands. General practitioners were encouraged to refer patients directly when arthritis was suspected. Patients were included when the symptom duration was less than 2 y. Patients from the BEST study were recruited between March 2000 and August 2002 at 20 centers in the western part of The Netherlands. Patients had a maximum disease duration of 2 y, were at least 18 y of age, and had active disease (defined as ≥ six swollen joints, ≥ six tender joints, and either an erythrocyte sedimentation rate of ≥ 28 mm/h or a global health assessment score of ≥ 20 on a 100-mm visual analog scale, where 0 = best and 100 = worst). Only patients with a diagnosis of RA were included in the present study. These cohorts are further described in detail in other reports [15,16].Sample set 3 consisted of 1,500 RA patients (70% RF-positive) and 1,000 unaffected participants (controls) from the Swedish EIRA study as previously described [17]. Briefly, RA patients and controls aged 18–70 y during May 1996 to December 2003 from a geographically defined area in the south and central regions of Sweden. Control participants were randomly selected from a continuously updated national population register, with consideration given to age, sex, and living area. If the selected control was not traceable, reported having RA, or refused to participate, a new control was selected using the same procedure.

Sample set 4, obtained from the Genomics Collaborative, (GCI), comprised 475 RF-positive RA patients and 475 individually matched controls from the US and has been described in detail elsewhere [5]. In brief, all case samples were from white North Americans who were RF-positive and whose condition met the 1987 American College of Rheumatology diagnostic criteria for RA. Control samples were taken from a pool of healthy white individuals with no medical history of RA. A single control was matched to each case on the basis of sex, age (±5 y), and ethnicity (grandparental country/region of origin).

Patients from sample sets 1 and 4 had considerably longer disease duration at inclusion (13.8 ± 10.1 y and 11.7 ± 10.0 y, respectively) as compared to patients from sample sets 2 and 3 (<3 y) (Table S1). All controls were healthy unrelated white individuals originating from the same geographical area as the patients. There was no overlap between cases and controls across all studies. All protocols and recruitments were approved by national and/or local institutional review boards, and informed written consent was obtained from all participants.

SNP Selection and Genotyping

We chose 40 polymorphisms spanning *TRAF1*/*C5* and their flanking genes *PHD finger protein 19* (*PHF19*) and *centrosomal protein 110 kDa* (*CEP110)* for this study (Table 1). We selected haplotype tagging SNPs (htSNPs) from the International HapMap Project database (http://www.hapmap.org/index.html) as well as random SNPs from the University of California Santa Cruz database (http://genome.ucsc.edu/) to ascertain maximum haplotype information for each of the genes and intergenic regions that are likely to harbor regulatory regions. Chromosomal locations of the SNPs were extracted from SNPPER (http://snpper.chip.org/) Goldenpath hg17, dbSNP build 123. Genotyping across all studies was performed as described in detail in Protocol S1. Three of the 40 SNPs were excluded from further analysis after reviewing results from sample set 1, two of which (SNP13/rs4837803 and SNP27/rs10119768) were not polymorphic and one (SNP35/rs7856420) of which deviated from Hardy-Weinberg equilibrium.

Statistical Analysis

Single SNP and haplotype analysis

Single SNP analysis and genetic model assessment was initially performed using SPSS version 12.0 (SPSS, http://www.spss.com/) in sample set 1. We did not find evidence of a recessive model for any SNPs. Since all SNPs were in Hardy-Weinberg equilibrium and adhered to the additive model of association, we performed further tests using allelic comparisons. Single- and multilocus allelic analyses were performed using Haploview version 3.32 (MIT, http://www.broad.mit.edu/mpg/haploview/) [18] with 40 SNPs in sample set 1 followed by six significant SNPs in sample set 2, the three most significant SNPs in sample set 3, and the single best-associating SNP in sample set 4. Odds ratios were calculated using Epi Info v6 (CDC, http://www.cdc.gov/epiinfo/). All *p*-values reported were two-sided. A p-value < 0.05 was considered significant. Based on the linkage disequilibrium (LD) structure in sample set 1, haplotype blocks were inferred under the algorithm of Gabriel et al. [19] in Haploview 3.32. To further minimize haplotypic uncertainty, we used the software TagSNPs version 1.0 (http://www-rcf.usc.edu/~stram/tagSNPs.html) to identify eight htSNPs from block 2 and ten htSNPs from block 3 (global $R^2_h > 98\%$) [20]. The R^2_h coefficient is the squared correlation between the true haplotype count (number of copies of a haplotype) and the haplotype count predicted by TagSNPs. We chose htSNPs so that haplotypes were predicted with a global R^2_h value of 0.95 or above, indicating a high accuracy. Haplotype analyses using these htSNPs were performed in Haploview 3.32. Global *p*-values for haplotype associations were calculated using the software Haplo.Stats version 1.2 (http://mayoresearch.mayo.edu/mayo) used for estimating haplotype effects under the generalized linear model [21]. htSNPs from blocks 2 and 3 in sample set 1 were further investigated by sliding-window analysis in Haploview 3.32 to determine the basis of the associated haplotypes.

Combining datasets

Odds ratios (OR)s from all sample sets were combined by the random model of the Cochran-Mantel-Haenszel test as implemented in EasyMA [22]. A Breslow-Day test of between-stage heterogeneity was also performed in EasyMA to test for consistency across sample sets [23]. We observed evidence of heterogeneity for SNP rs4836834 at $p < 0.05$ (Table S2).

Logistics regression

Forward conditional logistics regression was performed using all six significant SNPs from sets 1 and 2 in SPSS 12.0. Genotypes were coded as categorical variables 0, 1, and 2 with the nonassociated genotype as reference. No evidence of a recessive effect was observed, as detailed in Table S3.

Population attributable risk

Population-attributable risk (PAR) was calculated using the population prevalence of the exposure (P) and the relative risk associated with the exposure (R), as follows [24].

$$\mathrm{PAR} = \frac{P(R-1)}{1+P(R-1)}$$

Autoantibodies

Baseline laboratory parameters included IgM-RF (ELISA) and IgG-ACPA (ELISA, Immunoscan RA Mark2, Euro-Diagnostica, http://www.eurodiagnostica.com/). The cut-off for antibody positivity was set according to the manufacturer's instructions. Since baseline ACPA status was not available for most samples from the BEST study, we restricted our analysis to the EAC and EIRA cohorts. Autoantibody (RF and ACPA) and genotype status were available from 419 patients from the EAC (we additionally genotyped RF-negative and ACPA-negative patients in this cohort) and 1,395 patients from the EIRA study.

Severity

Radiographs of hands and feet were scored at baseline and 2 y (n = 278) using the Sharp–van der Heijde method (erosions and joint space narrowing of hands and feet) [25]. Data was available from 193 A carriers (AA+AG) and 85 non-A carriers (GG) from the EAC cohort. The prevalence of either RF (A carriers 0.9%, non-A carriers 56.5%, p = 0.485) or anti-citrullinated protein antibodies (ACPA) (A carriers 60.7%, non-A carriers 60.8%, p = 0.986) is not significantly different between the two groups analyzed. Differences in means between groups were calculated using sharp scores adjusted for baseline with a two-sided nonparametric Mann-Whitney test.

Transcription Factor Binding Sites

Transcription factor binding sites (TFBSs) were predicted using MAPPER (http://bio.chip.org/mapper/mapper-main), a latform that combines information from two well-known TFBS databases, TRANSFAC and JASPAR. The prediction is generated from a hidden Markov model and is based on experimentally determined binding sites in multiple genomes [26].

Results

The TRAF1/C5 Region Associates with RA

To investigate whether the *TRAF1*/*C5* region on Chromosome 9q33-34 associates with RA, we selected a total of 40 polymorphisms spanning these candidates and their flanking genes for genotyping. Tagging SNPs as well as random SNPs were included to ascertain maximum haplotype information for each of the genes and to ensure coverage of intergenic regions which may harbor regulatory polymorphisms.

Single SNP analysis performed in the first set of RA patients (n = 290) and controls (n = 254) revealed significant association between six SNPs in the *TRAF1*/*C5* region (SNPs 4, 7, 10, 14, 15, and 16) and RA (p = 0.0104, 0.0153, 0.0080, 0.0039, 0.0039, and 0.0250, respectively) (Table 1). One SNP in *PHF19* (SNP3/rs10985070) and one in *CEP110* (SNP36/rs10818503) also showed moderate association with RA (p = 0.0387 and 0.0257, respectively). None of the SNPs investigated showed evidence of a recessive mode of association.

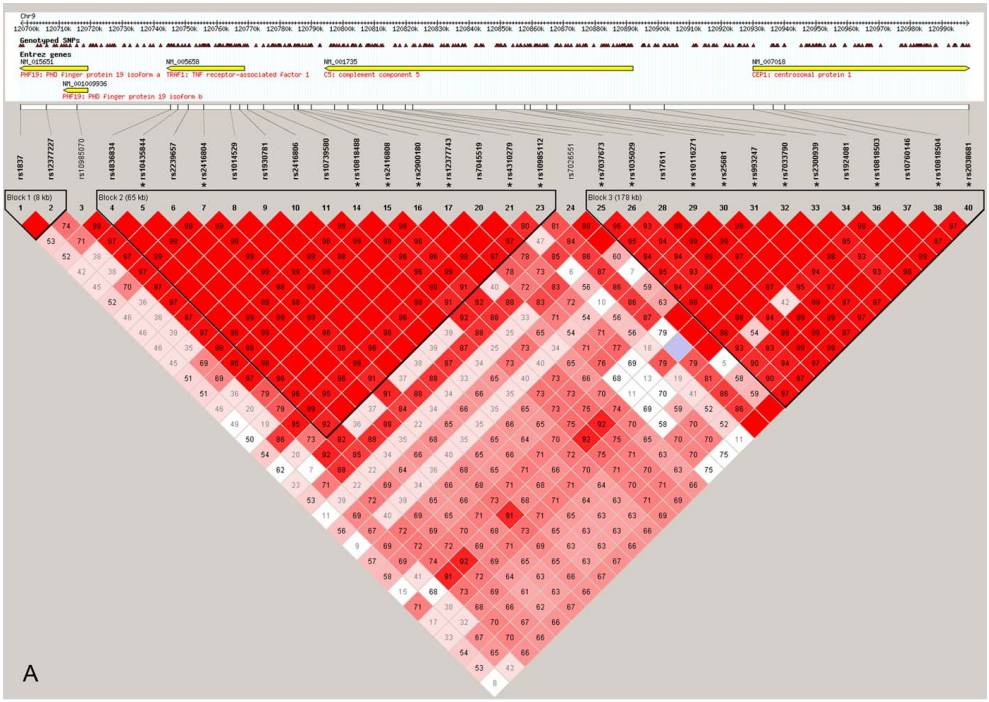
To delineate the haplotypic architecture, we estimated the underlying haplotype block structure of the 32 SNPs with minor allele frequency > 5% in the controls. We identified a potential recombination spot at SNP3/rs10985070 and SNP24/rs7026551 that divides the region into three inheritance blocks, blocks 1, 2, and 3. Block 1 is 8 kb and encompasses the 3′ end to intron 9 of *PHF19;* block 2 extends from *TRAF1* through 24 kb of intergenic sequence to exon 32 of *C5;* while block 3 (178 kb) contains the remainder of the *C5* gene and *CEP110* (Figure 1A). The LD structure in the Dutch population was similar to the structure reported by the International HapMap Project (unpublished data).

Table 1. Association of Candidate Gene SNPs with RA

SNP ID	SNP rs ID	Gene	Location	Chromosome Location[a]	Allele1/ Allele2	Susceptibility Allele	Allele Ratios[b]: Cases, Controls	Frequency[c]: Cases, Controls	Allelic OR (95% CI)	*p*-Value
1	rs1837	*PHF19*	3'UTR/intergenic	120697783	G/A	A	159:385, 110:290	0.292, 0.275	1.09 (0.81–1.47)	0.5611
2	rs12377227	*PHF19*	Intron 9	120706087	A/G	G	74:452, 44:354	0.141, 0.111	1.32 (0.87–2.00)	0.1742
3	rs10985070	*PHF19*	Intron 4	120715675	A/C	C	263:281, 168:236	0.483, 0.416	1.31 (1.01–1.72)	**0.0387**
4	rs4836834	*TRAF1*	Exon 8/3'UTR	120745455	A/T	T	260:268, 209:297	0.492, 0.413	1.38 (1.07–1.78)	**0.0104**
5	rs10435844	*TRAF1*	Intron 7	120747753	T/G	G	182:304, 158:326	0.374, 0.326	1.24 (0.94–1.62)	0.1169
6	rs2239657	*TRAF1*	Intron 7	120751074	A/G	G	207:329, 164:336	0.386, 0.328	1.29 (0.99–1.68)	0.0509
7	rs2416804	*TRAF1*	Intron 4	120755950	C/G	G	224:228, 206:288	0.496, 0.417	1.37 (1.05–1.79)	**0.0153**
8	rs1014529	*TRAF1*	Intron 3	120764497	G/C	C	208:332, 165:333	0.385, 0.331	1.26 (0.97–1.64)	0.0708
9	rs1930781	*TRAF1*	Intron 2	120767388	A/G	G	208:332, 166:336	0.385, 0.331	1.27 (0.98–1.65)	0.0668
10	rs2416806	*TRAF1*	5'UTR/intergenic	120769846	C/G	G	214:308, 160:326	0.410, 0.329	1.42 (1.09–1.85)	**0.0080**
11	rs10739580	*TRAF1*	5'UTR/intergenic	120774836	T/C	C	192:314, 154:320	0.379, 0.325	1.27 (0.97–1.67)	0.0741
12	rs6478491	*TRAF1*	5'UTR/intergenic	120778044	G/C	C	8:532, 7:493	0.015, 0.014	1.06 (0.35–3.26)	0.9123
14	rs10818488	*TRAF1*	5'UTR/intergenic	120784641	G/A	A	247:255, 193:289	0.492, 0.400	1.45 (1.12–1.88)	**0.0039**
15	rs2416808	*TRAF1*	5'UTR/intergenic	120785837	A/G	G	247:255, 193:289	0.492, 0.400	1.45 (1.12–1.88)	**0.0039**
16	rs2900180	*TRAF1*	5'UTR/intergenic	120785936	C/T	T	216:332, 161:331	0.394, 0.327	1.34 (1.03–1.74)	**0.0250**
17	rs12377743	*C5*	3'UTR/intergenic	120790004	G/A	A	58:482, 39:461	0.107, 0.078	1.42 (0.91–2.22)	0.1032
18	rs7874632	*C5*	3'UTR/intergenic	120792886	G/C	C	4:512, 2:500	0.008, 0.004	1.95 (0.31–15.38)	0.4323
19	rs12005198	*C5*	3'UTR/intergenic	120793949	A/G	A	539:1, 501:1	0.998, 0.998	1.08 (0.00–39.38)	0.9588
20	rs7045519	*C5*	Intron 39	120797394	C/A	A	58:482, 40:462	0.107, 0.080	1.39 (0.89–2.17)	0.1255
21	rs4310279	*C5*	Intron 38	120801702	A/G	G	148:392, 124:374	0.274, 0.249	1.11 (0.84–1.48)	0.3587
22	rs16910233	*C5*	Intron 37	120803165	C/G	G	4:518, 3:499	0.008, 0.006	1.28 (0.24–7.24)	0.7433
23	rs10985112	*C5*	Intron 32	120810962	G/A	A	54:486, 33:465	0.100, 0.066	1.57 (0.98–2.52)	0.0501
24	rs7026551	*C5*	Intron 31	120812687	A/C	C	121:393, 94:402	0.235, 0.190	1.32 (0.96–1.80)	0.0749
25	rs7037673	*C5*	Intron 28	120820038	C/T	C	319:195, 279:219	0.621, 0.560	1.28 (0.99–1.66)	0.0508
26	rs1035029	*C5*	Intron 27	120822372	A/G	G	206:316, 188:316	0.395, 0.373	1.10 (0.85–1.42)	0.4766
28	rs17611	*C5*	Exon 19	120848754	G/A	G	261:187, 267:227	0.583, 0.540	1.19 (0.91–1.55)	0.1935
29	rs10116271	*C5*	Intron 15	120857702	C/T	C	255:251, 221:263	0.504, 0.457	1.21 (0.93–1.56)	0.1362
30	rs25681	*C5*	Exon 13	120859559	C/T	C	297:205, 265:227	0.592, 0.539	1.24 (0.96–1.61)	0.0918
31	rs993247	*C5*	Intron 10	120864803	A/G	A	307:213, 274:226	0.590, 0.548	1.19 (0.92–1.54)	0.1717
32	rs7033790	*C5*	Intron 8	120867946	C/T	T	129:389, 100:402	0.249, 0.199	1.33 (0.98–1.81)	0.0565
33	rs2300939	*C5*	Intron 1	120891271	C/T	T	44:454, 34:452	0.088, 0.070	1.29 (0.79–2.11)	0.2856
34	rs1924081	*C5*	5'UTR/intergenic	120902001	T/A	A	137:413, 100:398	0.249, 0.201	1.32 (0.98–1.79)	0.0620
36	rs10818503	*CEP110*	Exon 1	120930324	A/G	G	181:333, 144:358	0.352, 0.287	1.35 (1.03–1.78)	**0.0257**
37	rs10760146	*CEP110*	Intron 2	120936639	C/T	C	332:216, 276:220	0.606, 0.556	1.23 (0.95–1.58)	0.1061
38	rs10818504	*CEP110*	Exon 5	120940243	C/T	C	312:210, 280:226	0.598, 0.553	1.20 (0.93–1.55)	0.1504
39	rs17292952	*CEP110*	Exon 16	120983394	G/A	G	525:25, 473:25	0.955, 0.950	1.11 (0.61–2.03)	0.7189
40	rs2038681	*CEP110*	Intron 25	120998594	A/C	C	94:416, 74:428	0.184, 0.147	1.31 (0.92–1.85)	0.1147

Data in bold indicate significant SNPs.

a. Chromosomal locations were extracted from SNPPER Goldenpath hg17, dbSNP build 123.
b. Number of alleles were compared in cases versus controls: allele1:allele2 cases, allele1:allele2 controls.
c. Frequency of the susceptibility allele.



B

Haplotype	Case,Control[¥]	Case, Control Frequency	OR (95%CI)	Global P	P
Block 1					
GA	309.6 : 236.4, 248.2 : 155.8	0.567, 0.614	0.82 (0.62 - 1.07)		0.1444
AA	159.3 : 386.7, 111.3 : 292.7	0.292, 0.276	1.08 (0.81 - 1.46)	0.212	0.5822
GG	77.0 : 469.0, 44.5 : 359.5	0.141, 0.110	1.31 (0.87 - 1.98)		0.1592
Block 2					
TCGACGAG	256.6 : 259.4, 291.4 : 214.6	0.497, 0.576	0.73 (0.57 - 0.95)		**0.0119**[a]
GGAGTGGG	134.8 : 381.2, 122.5 : 383.5	0.261, 0.242	1.11 (0.83 - 1.48)	0.047	0.4828
GGAGTGAG	62.9 : 453.1, 41.9 : 464.1	0.122, 0.083	1.54 (1.00 - 2.37)		**0.0393**[b]
TGAGCAAA	47.6 : 468.4, 32.6 : 473.4	0.092, 0.064	1.47 (0.91 - 2.39)		0.0984
Block 3					
TATTGCCATA	187.9 : 332.0, 200.1 : 298.1	0.361, 0.402	0.84 (0.65 - 1.10)		0.1866
CGCCATCGCC	93.5 : 426.4, 70.6 : 427.6	0.180, 0.142	1.33 (0.94 - 1.89)		0.0984
CGCCACCACA	66.6 : 453.4, 77.1 : 421.1	0.128, 0.155	0.81 (0.56 - 1.17)		0.2205
CACCACCACA	58.1 : 461.8, 43.7 : 454.5	0.112, 0.088	1.30 (0.84 -2.00)	0.106	0.2008
CATCACTGCA	44.0 : 476.0, 32.0 : 466.3	0.085, 0.064	1.35 (0.82 - 2.22)		0.2156
CGCCATCGCA	33.2 : 486.7, 26.1 : 472.1	0.064, 0.052	1.23 (0.70 - 2.16)		0.4348
CATTGCCATA	14.5 : 505.5, 17.3 : 480.9	0.028, 0.035	0.84 (0.39 - 1.79)		0.5269

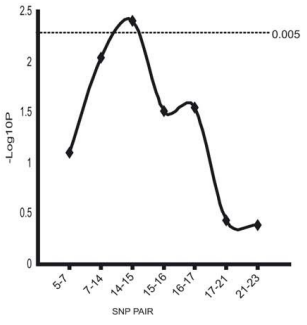


Figure 1. LD Structure and Haplotype Association across the TRAF1/C5 Region in Sample Set 1 (290 RA Patients and 254 Controls)

(A) Haplotype block structure was predicted on the basis of the strength of pairwise LD, which is presented as a 2×2 matrix; red represents very high LD. *(D'),* white indicates absence of correlation between SNPs, and blue indicates high correlation with a low level of significance. SNPs that were chosen for haplotype analysis are indicated along the top by an asterisk.

(B) Using htSNPs from each block, indicated by the asterisk in (A), haplotypes were inferred with a certainty of above 98% as represented by the R^2_h value. [¥]Comparisons were made between one haplotype versus all others in cases and controls. [a]The protective haplotype that is significantly less frequent in cases. [b]The susceptible haplotype that is significantly more frequent in patients.

(C) Sliding window of the susceptible haplotype using consecutive two-SNP combinations of the htSNPs reveals that SNP14 and SNP15 account for most of the association observed. For each of the SNP pairs we show the $-\log_{10}$ *p*-value. The dotted line indicates a nominal *p*-value of 0.005.

To further minimize haplotypic uncertainty and to identify the minimal combination of SNPs that provide maximum information content within each block, we scanned these two blocks independently using the software TagSNPs. In block 2 we identified a minimal set of eight htSNPs with a global R^2_h of 0.996, and in block 3 we found ten htSNPs with an R^2_h of 0.985, indicating that haplotypes can be inferred with >95% certainty. Haplotypes were predicted, and analyses from all blocks revealed that the association with RA was restricted to SNPs in block 2 (Figure 1B), as demonstrated by the global *p*-value of association ($p < 0.05$) and suggesting the possible involvement of *TRAF1* and/or the 3′ end of *C5*.

Of the four common haplotypes capturing > 95% of the variation, two significantly associated haplotypes were observed, one increased in RA patients (susceptible haplotypeb $p = 0.039$), and one over-represented in controls (protective haplotypea $p = 0.012$). By applying a two-marker sliding window analysis, we observed a significant peak centered on SNP14/rs10818488 and SNP15/rs2416808 ($p = 0.0039$) (Figure 1C). Using three- and five-marker windows did not alter the outcome, suggesting that the significance seen with the other *TRAF1* SNPs (Table 1) may be due to LD. To explore this possibility, we analyzed the r^2-values with respect to SNP14. We confirmed that the most significant SNPs, which are located in *TRAF1* and the adjacent intergenic region, are highly correlated with SNP14 ($r^2 > 0.90$) (Figure 2).

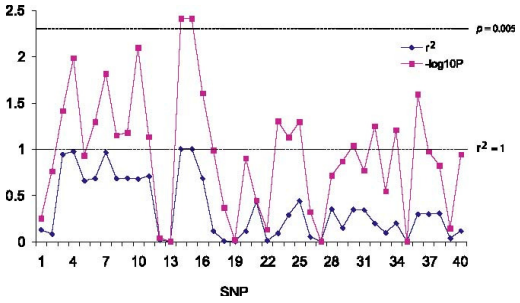


Figure 2. RA-Associated SNPs are Highly Linked to SNP14

Pairwise LD (r^2) between associated SNP14/rs10818488 and all other SNPs genotyped was calculated. For each of the SNPs listed along the x-axis we show the $-\log_{10}$ (y-axis) of the *p*-values for RA patients versus controls. Dotted lines indicate a nominal *p*-value of 0.005 and a maximal r^2 value of 1. In a logistics regression model, only SNP14/rs10818488 remained statistically significant ($p = 6.16 \times 10^{-4}$).

Replication in Three Independent Sample Sets from The Netherlands, Sweden, and the US

Six tagging SNPs which were significant ($p < 0.05$) in the initial study were selected for replication in a fully independent set of Dutch cases and controls (sample set 2). Of these, only SNP14/rs10818488 and SNP15/rs2416808 were statistically significant ($p < 0.05$). Haplotype and a two-marker sliding window analysis localized the strongest region of association to SNP14 and SNP15, confirming the results from sample set 1 (unpublished data). Combined analyses of the data from sample sets 1 and 2 showed an even stronger association for SNP14 (OR 1.34, 95% confidence interval [CI] 1.13–1.58; $p = 5.56 \times 10^{-4}$) and SNP15 (OR 1.33, 95% CI 1.13–1.57; $p = 6.65 \times 10^{-4}$) (Table 2). Although the other four SNPs did not reach statistical significance in sample set 2, they were highly significant in the combined analysis (Table 2). To evaluate putative modes of inheritance, we calculated genotype-specific ORs in the combined dataset as detailed in Table S3. All SNPs were consistent with an additive model. On the basis

of forward conditional logistics regression, SNP14 remained in the model with a heterozygote (AG) OR of 1.38 (95% CI 1.04–1.83, $p = 0.027$) and a homozygote (AA) OR of 2.06 (95% CI 1.42–2.98, $p = 1.29 \times 10^{-3}$).

Similar replication of three SNPs in a cohort of Swedish patients and controls (sample set 3) confirmed association with SNP14/rs10818488 ($p = 0.0078$) (Table 2). A combined analysis of patients and controls of European origin (Dutch and Swedish) with 2,244 RA patients and 1,524 controls (sample sets 1, 2, and 3) showed that the most significant associations could again be attributed to SNP14 (OR 1.24, 95% CI 1.11–1.38, $p = 1.73 \times 10-5$) and SNP15/rs2416808 (OR 1.23, 95% CI 1.09–1.40, $p = 7.21 \times 10^{-5}$) (Table 2). Additionally, these findings were further replicated in a case-control sample set from the US. Since LD analysis in the original three sample sets showed SNP14 and SNP15 to be highly correlated ($r^2 > 0.98$) we genotyped only SNP14 and confirmed that the minor susceptibility allele was associated with RA risk (OR 1.36, 95% CI 1.13–1.64; $p = 0.001$) (Table 2). Combined analysis of SNP14 in all four independent sets (($n_{cases/controls}$ = 2,719/1,999) yielded a highly significant association OR_{common}=1.26, 95% CI 1.15–1.37, $p_{combined} = 1.40 \times 10^{-8}$) and a PAR of 6.1% (95% CI 4.0–8.5).

Table 2. Association of Significant SNPs in Four Independent Sample Sets

SNP	Set 1 (Netherlands)			Set 2 (Netherlands)			Sets 1 + 2		
	Allele Ratios[a] Cases, Controls	Allelic OR (95% CI)	p-Value	Allele Ratios[a] Cases, Controls	Allelic OR (95% CI)	p-Value	Allele Ratios[a] Cases, Controls	Allelic OR (95% CI)[b]	p-Value
SNP4/rs4836834	260:268, 209:297	1.38 (1.07–1.78)	0.0104	378:432, 225:301	1.17 (0.93–1.47)	0.1626	638:700, 434:598	1.41 (0.96–2.13)	0.078
SNP7/rs2416804	224:228, 206:288	1.37 (1.05–1.79)	0.0153	369:415, 224:300	1.19 (0.95–1.50)	0.1242	593:643, 430:588	1.27 (1.07–1.50)	0.006
SNP10/rs2416806	214:308, 160:326	1.42 (1.09–1.85)	0.0080	296:489, 180:340	1.14 (0.90–1.45)	0.2559	510:798, 340:666	1.26 (1.02–1.56)	0.029
SNP14/rs10818488	247:255, 193:289	1.45 (1.12–1.88)	0.0039	387:425, 220:304	1.26 (1.00–1.58)	0.0419	634:680, 413:593	1.34 (1.13–1.58)	**5.56×10^{-4}**
SNP15/rs2416808	247:255, 193:289	1.45 (1.12–1.88)	0.0039	387:425, 221:303	1.25 (0.99–1.57)	0.0494	634:680, 414:592	1.33 (1.13–1.58)	**6.65×10^{-4}**
SNP36/rs10818503	181:333, 144:358	1.35 (1.03–1.78)	0.0257	259:527, 166:352	1.04 (0.82–1.33)	0.7329	440:860, 310:710	1.17 (0.98–1.40)	0.210

SNP	Set 3 (Sweden)			Sets 1 + 2 + 3		
	Allele Ratios[a] Cases, Controls	Allelic OR (95% CI)	p-Value	Allele Ratios[a] Cases, Controls	Allelic OR (95% CI)[b]	p-Value
SNP4/rs4836834	—	—	—	—	—	—
SNP7/rs2416804	—	—	—	—	—	—
SNP10/rs2416806	1,212:1,756, 677:1,089	1.11 (0.98–1.25)	0.0893	1,722:2,554, 1,017:1,755	1.18 (1.03–1.34)	0.015
SNP14/rs10818488	1,500:1,500, 840:984	1.17 (1.04–1.32)	**0.00781**	2,134:2,180, 1,253:1,577	1.23 (1.12–1.36)	**1.73×10^{-5}**
SNP15/rs2416808	1,351:1,353, 772:890	1.15 (1.02–1.30)	**0.0241**	1,985:2,033, 1,186:1,482	1.23 (1.09–1.40)	**7.21×10^{-5}**
SNP36/rs10818503	—	—	—	—	—	—

SNP	Set 4 (US)			Sets 1 + 2 + 3 + 4		
	Allele Ratios[a] Cases, Controls	Allelic OR (95% CI)	p-Value	Allele Ratios[a] Cases, Controls	Allelic OR (95% CI)[b]	p-Value
SNP4/rs4836834	—	—	—	—	—	—
SNP7/rs2416804	—	—	—	—	—	—
SNP10/rs2416806	—	—	—	—	—	—
SNP14/rs10818488	427:513, 356:582	1.36 (1.13–1.64)	**0.00102**	2,561:2,693, 1,609:2,159	1.26 (1.15–1.37)	**1.40×10^{-8}**
SNP15/rs2416808	—	—	—	—	—	—
SNP36/rs10818503	—	—	—	—	—	—

Data in bold indicate the p-value of the most significant SNPs.
a. The allele frequencies between allele1:allele2 in cases was compared to allele1:allele2 in controls. Allele 1 refers to the susceptibility allele from Table 1.
b. Mantel-Haenzel OR as calculated under the random model. Raw data can be obtained from Table S2.

Association with Autoantibody-Positive Disease

RA is a heterogeneous disease with a considerable variation in phenotype as evidenced by the fact that some patients are autoantibody-positive whereas others are not. Antibodies to citrullinated protein antigens, called ACPAs, have gained much interest as current data suggest

that ACPA-positive and negative RA may have different genetic risk factors [27]. To investigate whether the *TRAF1*/*C5* region is associated with a specific phenotype of RA, we next stratified patients for autoantibody status from whom baseline ACPA and RF measurements were available (n = 1,814) Interestingly, SNP14/rs10818488 mainly predisposes to autoantibody-positive disease when compared to controls (OR 1.25, 95% CI 1.11–1.40, $p = 2.27 \times 10^{-4}$) (Figure 3A). Although we also observed an increase in the frequency of the A allele in autoantibody-positive as compared to autoantibody-negative disease, this difference did not reach formal statistical significance (OR 1.15, 95% CI 0.98–1.34, p = 0.0789). These data therefore suggest that the current genetic risk factor may be predominant in the autoantibody-positive subset of RA patients.

Association with Severity

Because the clinical course of RA can vary considerably ranging from nonerosive disease to rapidly progressive joint damage, we also analyzed whether the SNPs in the *TRAF1/C5* region were involved with RA progression. Annual X-rays of the hands and feet of patients were assigned Sharp–van der Heijde units, a combined score for bone erosions and joint space narrowing. Carriers of the minor susceptibility allele of SNP14/rs10818488 had an almost 2-fold higher severe disease course at 2 y after inclusion as compared to the non-A carriers (Figure 3B; mean ± SE score of A carriers/non-A carriers, 11.4 ± 1.4/7.1 ± 1.8; p = 0.008) indicating that the A allele predisposes not only to RA susceptibility, but also to severity.

	Number of Alleles			Allele Frequency		Cases vs Controls		AutoAntibody Positive vs Negative	
	A	G	Total	A	G	Allelic OR (95% CI)	P	Allelic OR (95% CI)	P
Controls*	1253	1577	2830	0.44	0.56	-	-		
ACPA-RF-	482	558	1040	0.46	0.54	1.09 (0.94 - 1.26)	0.2509	-	-
ACPA+RF-	120	136	256	0.47	0.53	1.11 (0.85 - 1.45)	0.4228	1.02 (0.77 - 1.36)	0.8792
RF+ACPA-	216	224	440	0.49	0.51	1.21 (0.99 - 1.49)	0.0588	1.12 (0.89 - 1.40)	0.3336
RF+ACPA+	941	951	1892	0.50	0.50	1.25 (1.11 - 1.40)	**2.27×10^{-4}**	1.15 (0.98 - 1.34)	**0.0789**

* Controls refer to the combined controls from sets 1, 2 and 3.

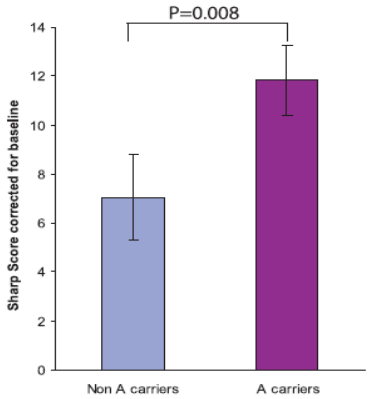


Figure 3. The A Allele of SNP14 Is Associated with Distinct Phenotypes of RA. (Top, table) The frequency of the A allele of patients from whom baseline autoantibody status (ACPA and RF) were available was calculated. ORs and pvalues were calculated between each subgroup and controls, and indicated a predominantly higher frequency of the A allele in ACPA- and RF-positive patients. (Bottom, bar graph) Progression of joint damage in Sharp–van der Heijde units (“Sharp score”) is higher in the presence of the minor A allele of SNP14. Radiological data of 193 A carriers and 85 non-A carriers were available, and differences between the groups were calculated based on disease severity after 2 y corrected for baseline.

SNP14/rs 10818488 Is Located in a TFBS

To investigate the potential functional effect of this SNP, we scanned for transcription factor binding sites using MAPPER [26]. The SNP14 susceptibility A allele encodes a potential binding site for EP300, a histone acetyl transferase that regulates transcription via chromatin remodeling. In the absence of this allele, the binding of EP300 to this region is predicted to be disrupted, potentially disturbing the epigenetic tag for transcriptional activation. We hypothesize that this putative transcription factor binding site may be involved in the regulation of the neighbouring *TRAF1* and/or *C5* gene.

Discussion

Using a candidate-gene approach, we identified the *TRAF1*/*C5* region on Chromosome 9q33-34 as a susceptibility and severity factor for RA. This region was also associated with RA in a large-scale genetic association study (Schrodi et al., unpublished data). It is, therefore, intriguing to see that these independent studies, in which the process leading to results differed, give similar results, and in doing so provide strong evidence for the *TRAF1/C5* region as a true RA-associated genetic variant. The recent genome-wide study performed by the Wellcome Trust failed to identify this region as a candidate for RA [28]. Although it is difficult to speculate why this region was not detected in the Wellcome Trust Case Control Consortium study, we do note that none of the SNPs showing strong association in our hands was genotyped by the Wellcome Trust. Additionally, in line with our finding that this genetic risk factor is predominant in the autoantibody-positive subgroup, substratifications of the specific RA phenotypes may be needed to detect significant association.

The protein encoded by *TRAF1* is a member of the TNF receptor-associated factor (TRAF) protein family, which associates with and mediates the signal transduction from various receptors of the TNF receptor superfamily, including the receptor for TNFα [29]. In addition to a direct role in TNFα signaling, *TRAF1* has also been implicated in the activation and proliferation of T cells [30] and is expressed ubiquitously by other cells of the immune system including monocytes and B cells [31]. It is therefore possible that TRAF1 could play a role in RA by aiding the maintenance of the proinflammatory environment. Likewise, studies have also shown that perpetuation of inflammation coincides with increased levels of the anaphylatoxin C5a in the synovial fluid of RA patients [9]. Further studies in mice identifying *C5* as a candidate gene and showing that *C5* deficiency results in lower incidence and less-severe disease course support the role of this gene in inflammation [7,11]. It is therefore likely that although the primary function of the complement system is to protect the host from microorganisms, a deregulated activity of its central component, *C5,* can play a substantial role in inflammatory diseases as well.

In order to capture the variation within these candidate genes and potential regulatory regions, we genotyped both SNPs that were intragenic and those located 5′ and 3′ of the genes. Interestingly, the strongest replicated association was observed with SNP14/rs10818488, which maps to an intergenic region ~10 kb from both *TRAF1* and *C5* and is present in a TFBS which may regulate the transcriptional activity of its neighbouring genes. However, formal testing of all known variation within this region, both genetic and biological, will be necessary to pinpoint the precise biological process that is altered by the RA-associated variant(s) present in this region.

We found a strong association of this region in all four independent sample sets which represent varying disease durations (<3 to >10 y) as well as a correlation with disease progression. More importantly, these phenotypic data on joint destruction not only indicate that the *TRAF1*/*C5* region predisposes to RA, but also suggest that within the RA population, patients harboring the minor susceptibility allele of SNP14/rs10818488 tend to experience a more severe disease course. Although the above findings most likely exclude the possibility of a spurious association, especially since each case group was assigned an ethnically and geographically matched control group, background levels of population stratification as described by Cardon et al. [32]

may exist in the different populations under study. It is therefore conceivable that the slight variation in the observed effect between the four populations may partially be due to varying sample sizes and background levels of population admixture. RA is a common complex disease that results from the interaction of multiple genetic variants, each with relatively low penetrance, with an array of environmental triggers [33]. In advance of a genetic profile that can accurately pinpoint individuals at risk, identification of these genetic variants can provide insight into the underlying mechanisms of disease and the specific pathways associated with disease induction and/or progression. Understanding the function of these common disease-associated variants will be important to identify potential targets for intervention strategies that could prove useful to all patients, whether or not they carry the disease-associated variant.

In summary, this study provides robust evidence from four independent sample sets (two of Dutch origin, one of Swedish descent, and one from the US) that genetic variants within the *TRAF1/C5* region are associated with RA, indicating a possible role for these immune-related genes in the biological process underlying RA disease pathogenesis.

Limitations of This Study

Our study defines the *TRAF1/C5* as a novel genetic region present in the human genome that predisposes to RA. However, the causative variation (SNP) or the biological mechanism explaining this association is not yet known. Although it could be that the current identified polymorphism is causative, other proxies in high LD with this SNP could also be responsible for this issue, which can be addressed in more detailed by functional studies. Furthermore, although our data indicate a predominant association with autoantibody-positive disease, our study is underpowered to exclude the possibility that the *TRAF1/C5* region also predisposes to autoantibody-negative disease. By combining information obtained from other cohorts in which both the autoantibody and the *TRAF1/C5* status are known, this question should be resolved in the future. As it has been indicated that distinct genetic risk factors underlie either autoantibody-positive or autoantibody-negative disease, such additional information would provide more detailed knowledge on the genetic heterogeneity of these two distinct phenotypes of RA.

References

(1) Firestein GS (2003) Evolving concepts of rheumatoid arthritis. Nature 423: 356–361.

(2) Seldin MF, Amos CI, Ward R, Gregersen PK (1999) The genetics revolutionand the assault on rheumatoid arthritis. Arthritis Rheum 42: 1071–1079.

(3) MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, et al. (2000) Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 43: 30–37.

(4) Deighton CM, Walker DJ, Griffiths ID, Roberts DF (1989) The contribution of HLA to rheumatoid arthritis. Clin Genet 36: 178–182.

(5) Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, et al. (2004) A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 75: 330–337.

(6) Maini RN (2004) Current and new antitumor necrosis factor agents in perspective. Arthritis Res Ther 6: S1–S2.

(7) Wang Y, Kristan J, Hao L, Lenkoski CS, Shen Y, et al. (2000) A role for complement in antibody-mediated inflammation: C5-deficient DBA/1 mice are resistant to collagen-induced arthritis. J Immunol 164: 4340–4347.

(8) Wang Y, Rollins SA, Madri JA, Matis LA (1995) Anti-C5 monoclonal antibody therapy prevents collagen-induced arthritis and ameliorates established disease. Proc Natl Acad Sci U S A 92: 8955–8959

(9) Hogasen K, Mollnes TE, Harboe M, Gotze O, Hammer HB, et al. (1995) Terminal complement pathway activation and low lysis inhibitors in rheumatoid arthritis synovial fluid. J Rheumatol 22: 24–28.

(10) Grant EP, Picarella D, Burwell T, Delaney T, Croci A, et al. (2002) Essential role for the C5a receptor in regulating the effector phase of synovial infiltration and joint destruction in experimental arthritis. J Exp Med 196: 1461–1471.

(11) Ji H, Gauguier D, Ohmura K, Gonzalez A, Duchatelle V, et al. (2001) Genetic influences on the end-stage effector phase of arthritis. J Exp Med 194: 321–330.

(12) Eyre S, Barton A, Shephard N, Hinks A, Brintnell W, et al. (2004) Investigation of susceptibility loci identified in the UK rheumatoid arthritis whole-genome scan in a further series of 217 UK affected sibling pairs. Arthritis Rheum 50: 729–735.

(13) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, et al. (1988) The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 31: 315–324.

(14) Brinkman BM, Huizinga TW, Kurban SS, van der Velde EA, Schreuder GM, et al. (1997) Tumour necrosis factor alpha gene polymorphisms in rheumatoid arthritis: association with susceptibility to, or severity of, disease? Br J Rheumatol 36: 516–521.

(15) Goekoop-Ruiterman YP, Vries-Bouwstra JK, Allaart CF, van Zeben D, Kerstens PJ, et al. (2005) Clinical and radiographic outcomes of four different treatment strategies in patients with early rheumatoid arthritis (the BeSt study): A randomized, controlled trial. Arthritis Rheum 52: 3381–3390.

(16) van Aken J, van Bilsen JH, Allaart CF, Huizinga TW, Breedveld FC (2003) The Leiden Early Arthritis Clinic. Clin Exp Rheumatol 21: S100–S105.

(17) Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L (2004) A geneenvironment interaction between smoking and shared epitope genes in HLA-DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 50: 3085–3092.

(18) Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.

(19) Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, et al. (2002) The structure of haplotype blocks in the human genome. Science 296: 2225–2229.

(20) Stram DO, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, et al. (2003) Choosing haplotype-tagging SNPS based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. Hum Hered 55: 27–36.

(21) Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. Am J Hum Genet 70: 425–434.

(22) Cucherat M, Boissel JP, Leizorovicz A, Haugh MC (1997) EasyMA: A program for the meta-analysis of clinical trials. Comput Methods Programs Biomed 53: 187–190.

(23) Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN (2003) Metaanalysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat Genet 33: 177–182.

(24) Breslow NE, Day NE (1980) Statistical methods in cancer research. Volume I—The analysis of case-control studies. IARC Sci Publ 5–338.

(25) van der Heijde DM (1996) Plain X-rays in rheumatoid arthritis: Overview of scoring methods, their reliability and applicability. Baillieres Clin Rheumatol 10: 435–453.

(26) Marinescu VD, Kohane IS, Riva A (2005) MAPPER: A search engine for the computational identification of putative transcription factor binding sites in multiple genomes. BMC Bioinformatics 6: 79.

(27) van Gaalen F, Ioan-Facsinay A, Huizinga TW, Toes RE (2005) The devil in the details: The emerging role of anticitrulline autoimmunity in rheumatoid arthritis. J Immunol 175: 5575–5580.

(28) The Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447: 661–678.

(29) Wajant H, Henkler F, Scheurich P (2001) The TNF-receptor-associated factor family: Scaffold molecules for cytokine receptors, kinases and their regulators. Cell Signal 13: 389–400.

(30) Sabbagh L, Srokowski CC, Pulle G, Snell LM, Sedgmen BJ, et al. (2006) A critical role for TNF receptor-associated factor 1 and Bim down-regulation in CD8 memory T cell survival. Proc Natl Acad Sci U S A 103: 18703–18708.

(31) Zapata JM, Krajewska M, Krajewski S, Kitada S, Welsh K, et al. (2000) TNFRassociated factor family protein expression in normal tissues and lymphoid malignancies. J Immunol 165: 5084–5096.

(32) Cardon LR, Palmer LJ (2003) Population stratification and spurious allelic association. Lancet 361: 598–604.

(33) Kallberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, et al. (2007) Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. Am J Hum Genet 80: 867–875.

Supporting Information

Protocol S1. Genotyping Methods
Found at doi:10.1371/journal.pmed.0040278.sd001 (26 KB DOC).

Table S1. Clinical Characteristics of RA Patients from Four Independent Sample Sets
Found at doi:10.1371/journal.pmed.0040278.st001 (21 KB XLS).

Table S2. Mantel-Haenzel OR and P for Combined Sets
Found at doi:10.1371/journal.pmed.0040278.st002 (22 KB XLS).

Table S3. Genotypic Odds Ratios for the Most Significant SNPs in Sample Sets 1 and 2
Found at doi:10.1371/journal.pmed.0040278.st003 (21 KB XLS).

Table S4. Primers and Conditions for PCR_RFLP
Found at doi:10.1371/journal.pmed.0040278.st004 (21 KB XLS).

Accession Numbers

The GenBank (http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed) accession numbers for the genes discussed in this article are C5 (NM_001735); CEP110 (NM_007018); PHF19 (XM_045308); TRAF1 (NM_005658). The ExPASy UniProtKB/Swiss-Prot (http://www.expasy.org/sprot/) accession number for EP300 is Q09472.

Chapter 3

Replication of the *Tumor Necrosis Factor Receptor Associated Factor 1/Complement Component* 5 Region as a Susceptibility Locus for Rheumatoid Arthritis in a European Family-Based Study

F.A.S. Kurreeman
D. Rocha
J. Houwing-Duistermaat
S. Vrijmoet
V. H. Teixeira
P. Migliorini,
A. Balsa
R. Westhovens
P. Barrera
H. Alves
C. Vaz
M. Fernandes
D. Pascual-Salcedo,
L. Michou
S. Bombardieri
T. Radstake
P. van Riel
L. van de Putte
A. Lopes-Vaz
B. Prum
T. Bardin
I. Gut
F. Cornelis
T.W.J. Huizinga
E. Petit-Teixeira
R.E.M. Toes

European Consortium on Rheumatoid Arthritis Families

Arthritis Rheum. 58, 2670-2674 (2008).

Abstract

Objective

We recently showed, using a candidate gene approach in a case-control association study, that a 65-kb block encompassing *tumor necrosis factor receptor associated factor 1* (*TRAF1*) and Complement component 5 (*C5)* is strongly associated with rheumatoid arthritis (RA). Compared with case-control association studies, family-based studies have the added advantage of controlling potential differences in population structure and are not likely to be hampered by variation in population allele frequencies, as is seen for many genetic polymorphisms, including the *TRAF1/C5* locus. The aim of this study was to confirm this association in populations of European origin by using a family-based approach.

Methods

A total of 1,356 western European white individuals from 452 "trio" families were genotyped for the rs10818488 polymorphism, using the TaqMan allelic discrimination assay.

Results

We observed evidence for association, demonstrating departure from Mendel's law, with an overtransmission of the rs10818488 A allele (A = 55%; P = 0.036). By taking into consideration parental phenotypes, we also observed an increased A allele frequency in affected versus unaffected parents (A = 64%; combined *P* = 0.015). Individuals carrying the A allele had a 1.2-fold increased risk of developing RA (allelic odds ratio 1.24, 95% confidence interval 1.04–1.50).

Conclusion

Using a family-based study that is robust against population stratification, we provide evidence for the association of the *TRAF1/C5* rs10818488 A allele and RA in populations of European descent, further substantiating our previous findings. Future functional studies should yield insight into the biologic relevance of this locus to the pathways involved in RA.

Introduction

Rheumatoid arthritis (RA) is one of the most common autoimmune diseases, affecting ~1% of the population worldwide. Environmental as well as genetic factors are thought to play an important role in both the onset and the progression of the disease (1). Because the genetic contribution to RA has been estimated to be 50-60%, the identification of genes contributing to the disease is important for the understanding of underlying biologic mechanisms (2).

In addition to HLA, the first identified genetic risk factor (3), and 3 other replicated regions including the protein tyrosine phosphatase N22 gene (4), the 6q23 locus near the *tumor necrosis factor (TNF)α induced protein 3* gene (5,6), and the STAT4 gene (7,8), we recently reported a new genetic locus associated with RA (9). This region encompasses the *TNF receptor associated factor 1 gene (TRAF1*) as well as the *complement component factor 5 (C5)* gene, both of which are immune regulators and potential perpetuators of inflammation. We identified 1 single-nucleotide polymorphism (SNP), rs10818488, located within a 65kb haplotype block, explaining most of the association signal in this region in several populations, including Dutch, Swedish, and US sample sets consisting of 2,719 patients with RA and 1,999 control subjects (odds ratio [OR] 1.28, 95% confidence interval [95% CI] 1.17-1.39, $P = 1.40 \times 10^{-8}$). Interestingly, the same genetic association was described in a recent whole-genome association study (10).

Our data revealed that although the case-control allele frequency increase in different sample sets ranged from 4% to 9%, the population frequency ranged from 38% to 46% in populations of European ancestry. Given that association studies compare the frequencies in patients versus healthy individuals, unknown biases in control frequencies may lead to spurious associations. Thus, family-based association studies remain important to definitively establish association, especially when small effect sizes as well as variability in allele frequency in different populations are observed. Therefore, with the aim to further substantiate the association at the *TRAF1/C5* locus, we took advantage of one of the largest reported European family resources dedicated to RA family-based studies.

Patients & Methods

Study population

DNA was available from 452 white trio families from western Europe, through the European Consortium on Rheumatoid Arthritis Families. Each family consisted of 1 patient with RA and both of his or her parents. Ethnicity was determined by the origin of the grandparents. At the time of inclusion in the study, all patients with RA fulfilled the American College of Rheumatology (formerly, the American Rheumatism Association) 1987 revised criteria (11). All individuals provided written informed consent, and the study was approved by the ethics committees in each country.

For each patient, the characteristics collected were sex, age at the onset of RA, disease duration (years), presence of bone erosions on radiographs, presence of rheumatoid nodules, and seropositivity for rheumatoid factor (RF). RF status was not available for 9 of the patients with RA. Because the anti-citrullinated protein antibody (ACPA) status was available for only a small proportion of the patients (n 197), we did not perform further analyses for ACPAs. The 452 families included 313 families from France, 53 from Italy, 37 from Spain, 22 from Belgium, 13 from The Netherlands, and 14 from Portugal. The characteristics of the French and European sample sets are summarized in Table 1. None of the cases reported in this study overlap with our previous case-control study (9).

Genotyping

All DNA samples were genotyped using the TaqMan allelic discrimination assay according to the manufacturer's instructions (Applied Biosystems, Foster City, CA). At least 2 positive controls and 1 negative control were performed in each plate, and no inconsistencies were detected. Furthermore, the concordance rate between 10% random samples genotyped in duplicate was 99%. The genotyping success rate in the 452 trio families was 100% (parents and probands included).

Statistical analysis

The family-based analysis was performed using the basic transmission disequilibrium test (TDT) combined with parental phenotype information (parenTDT), as implemented in Haploview 4.0 (12). The basic TDT compares, for a given allele, the transmission of that allele from heterozygous parents to a child with RA, to the expected 50% transmission according to Mendel's law. The parental discordance test is based on counting the number of alleles in affected versus unaffected parents, treating each nuclear family parental pair as a matched pair. These counts combined with the transmitted and untransmitted counts of the basic TDT give a combined test statistic (13).

The current data set included 42 families with 1 affected parent. The control genotypes were derived from the untransmitted parental chromosomes, using Unphased version 3.0.10 (14). By combining the case-control data from the probands with parental data, genotypic OR were calculated using a conditional logistic regression model stratifying on matched pairs (each proband-pseudocontrol as a matched pair; each affected and unaffected parent as a matched pair). Robust standard errors were computed, taking into account the dependency between pairs from the same family. These analyses were performed using Stata version 9.0 software (www.stata.com).

Table 1. Characteristics of the patients with rheumatoid arthritis (RA)

Characteristic	All (n 452)	French (n 313)	European (n 139)
Female sex, no. (%)	393 (86.9)	275 (87.8)	118 (84.9)
Age at onset of RA, mean SD years	30.8 9.4	30.7 9.4	30.9 9.4
Disease duration, mean SD years	10.3 7.9	12.0 8.1	6.6 5.9
Bone erosions, no. (%)	340 (75.2)	246 (78.6)	94 (67.6)
Rheumatoid factor positivity, no. (%)	321 (71.0)	221 (70.6)	100 (71.9)
Nodules, no. (%)	72 (15.9)	55 (17.6)	17 (12.2)

Results

A total of 1,356 European individuals from 452 trio families (1 patient with RA and both parents) were analyzed. Three hundred thirteen families were of French origin, and 139 were from other continental western European countries (Table 1). No deviation from Hardy-Weinberg equilibrium was detected in parental genotypes (n = 904; P = 0.548). We observed deviation from Mendel's first law, with a 55% overtransmission of the A allele to the patients in the 452 families (P = 0.036) (Table 2) along with an increased prevalence of the A allele in affected parents (64% increase versus unaffected parents). By applying a parenTDT test that takes into consideration both the transmission from parents to patients and the occurrence of alleles in discordant parents (discordant for both disease status and genotype), we obtained a combined statistic (P = 0.015). There were no differences between families of French origin and those from other continental western European countries (55% and 56% overtransmission, respectively) as well as no differences between paternal and maternal transmissions (56% and 57% overtransmission, respectively).

One of the advantages of family trio data is that such data provide perfectly matched control subjects for each patient investigated. Each patient chromosome transmitted by a given parent is perfectly matched for the population of origin with the untransmitted chromosome of each heterozygous parent. We observed an rs10818488 A allele frequency of 36% in control subjects, increasing to 41% in the patients with RA (data not shown). Using conditional logistic regression for the combined data set of case-control and parental discordant pairs, we observed an allelic OR of 1.24 (95% CI 1.04-1.50) (Table 3).

Table 2. Family-based association of the TRAF1/C5 region with rheumatoid arthritis (RA)

Sample set	Trio families	rs10818488 A allele		Transmission†	*P*‡	Discordant parents§	*P*¶
		Transmitted	Untransmitted				
All	452	231	188	55	0.036	64	0.015
RF+	335	175	136	56	0.027	67	0.0099
RF−	108	47	49	49	0.92	40	0.84

* *TRAF1* = tumor necrosis factor receptor–associated factor 1; RF = rheumatoid factor.
† Percent transmission of the rs10818488 A allele from heterozygous parents.
‡ By standard transmission disequilibrium test (TDT), as implemented in Haploview 4.0.
§ Percent of the rs10818488 A allele from the affected parents in discordant parent pairs.
¶ Combined TDT with parental phenotype information statistic, as described by Purcell et al (13).

Table 3. Case–control association of *TRAF1/C5* rs10818488 with RA

Sample set	AA†	AG†	Allelic‡
All	1.48 (0.87–2.51)	1.37 (1.01–1.86)	1.24 (1.04–1.50)
RF+	1.93 (1.01–3.69)	1.47 (1.03–2.10)	1.31 (1.06–1.63)
RF−	0.56 (0.18–1.69)	1.16 (0.63–2.12)	1.00 (0.69–1.45)

* *TRAF1* _ tumor necrosis factor receptor–associated factor 1; RA = rheumatoid arthritis; RF = rheumatoid factor.
† Genotype-specific odds ratios (ORs) (95% confidence intervals [95% CIs]), using GG as the referent.
‡ Allelic ORs (95% CIs), using the G allele as the referent.

Because RA is a heterogeneous disease, and distinct subsets of patients are characterized by the presence of autoantibodies such as RF and ACPAs, we also performed a stratified analysis for the presence or absence of RF (only limited data were available for ACPAs [see Patients and Methods]). In concordance with our previous findings, we observed an overtransmission (A transmitted 56%) as well as a higher prevalence of the A allele in affected parents (A 67%; P = 0.0099, by parenTDT) in the 335 RF-positive families (Table 2).

Among RF-positive individuals, harbouring 1 copy of the risk allele yielded an OR of 1.47 (95% CI 1.03-2.10), and homozygous individuals had an almost 2-fold increased risk of developing RA (OR 1.93, 95% CI 1.01–3.69) (Table 3). Interestingly, this effect was not detected in the RF-negative subgroup, as reflected by transmission of 49% and an A allele frequency in affected parents of 40% in the 108 RF-negative families (*P* = 0.84) (Table 2), resulting in an allelic OR of 1.00 (95% CI 0.69–1.45) (Table 3).

Discussion

In the current family-based study, we observed evidence of association between RA and the *TRAF1/C5* rs10818488 A variant in a western European sample set, replicating our initial findings. One of the advantages of the family trio design is that it provides accurate estimations of matched control subjects for each patient, an approach that is robust against population stratification. In this set of perfectly matched cases and controls, we observed a 5% increase in the A allele frequency in the overall sample set and a 6% increase in the group of RF-positive patients of western European descent. Because these differences are well within the previously observed effect in the Dutch (9%), Swedish (4%), and American (7%) populations, these data together indicate that the contribution of the *TRAF1/C5* locus to RA is not likely caused by underlying stratification in populations of European descent, and that the effect size is modest (OR 1.3).

Although we observed association in the overall sample set as well as in the RF-positive subset of patients with RA, we did not observe any overtransmissions in the RF-negative subset of patients with RA (49% versus the expected 50%), indicating that the effect of this genetic risk factor in the overall sample set is likely to be attributable to the RF-positive subset of patients with RA. Given the recent finding of this locus by a genomewide SNP association study in ACPA-positive patients with RA, the current evidence further substantiates the role of the *TRAF1/C5* region in the autoantibodypositive subset of patients with RA. However, whether this association extends to the autoantibody-negative subset of patients remains to be determined, because our RF-negative sample set possesses only 26% power to detect small effect sizes (OR 1.3) at a significance level less than 0.05. Therefore, our data do not allow a conclusion regarding the (lack of) contribution of the *TRAF1/C5* locus to autoantibody-negative RA.

The currently identified polymorphism lies in an intergenic region between *TRAF1* and C5. Because *TRAF1* is involved in TNF-mediated signaling and C5 generates C5a, the most potent chemoattractant involved in inflammation, both genes possess characteristics relevant to the pathogenesis of RA. Linkage disequilibrium patterns have so far revealed that the haplotype block surrounding this polymorphism encompasses the *TRAF1* gene as well as the 3region of C5. Interestingly, our group recently observed an association of this SNP with the polyarticular form of juvenile idiopathic arthritis (15). It is therefore plausible that the association of the *TRAF1/C5* variant may not be restricted to RA as such but may be relevant for other diseases. Furthermore, because current HapMap phase II data suggest that the frequency of the rs10818488 A allele is high in Chinese, Japanese, and Yoruban populations (44±69%), it would be interesting to investigate whether this genetic risk factor is also relevant across various ethnic populations.

In conclusion, this study provides evidence of the association of the *TRAF1/C5* locus as one of the widely confirmed genetic risk factors for RA in white individuals of European descent. Endeavors to characterize the functional relevance of this polymorphism and/or others highly linked to it will yield insights into the biologic effects of this locus and will generate further crucial information on the pathways underlying disease.

References

(1) Firestein GS. Evolving concepts of rheumatoid arthritis [review]. Nature 2003;423:356–61.

(2) Van der Helm-van Mil AH, Wesoly JZ, Huizinga TW. Understanding the genetic contribution to rheumatoid arthritis. Curr. Opin Rheumatol 2005;17:299–304.

(3) Huizinga TW, Amos CI, van der Helm-van Mil AH, Chen W, van Gaalen FA, Jawaheer D, et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA–DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005;52:3433–8.

(4) Michou L, Lasbleiz S, Rat AC, Migliorini P, Balsa A, Westhovens R, et al. Linkage proof for PTPN22, a rheumatoid arthritis susceptibility gene and a human autoimmunity gene. Proc Natl Acad Sci U S A 2007;104:1649–54.

(5) Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, et al. Rheumatoid arthritis association at 6q23. Nat Genet 2007;39:1431–3.

(6) Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007;39:1477–82.

(7) Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, Gregersen PK. Association of STAT4 with rheumatoid arthritis in the Korean population. Mol Med 2007;13:455–60.

(8) Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357:977–86.

(9) Kurreeman FA, Padyukov L, Marques RB, Schrodi SJ, Seddighzadeh M, Stoeken-Rijsbergen G, et al. A candidate gene approach identifies the *TRAF1/C5* region as a risk factor for rheumatoid arthritis. PLoS Med 2007;4:e278.

(10) Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis: a genomewide study. N Engl J Med 2007;357:1199–209.

(11) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

(12) Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21:263–5.

(13) Purcell S, Sham P, Daly MJ. Parental phenotypes in family-based association analysis [review]. Am J Hum Genet 2005;76:249–59.

(14) Dudbridge F. Pedigree disequilibrium tests for multilocus haplotypes. Genet Epidemiol 2003;25:115–21.

(15) Albers HM, Kurreeman FA, Houwing-Duistermaat JJ, Brinkman DM, Kamphuis SS, Girschick HJ, et al. The *TRAF1/C5* region is a risk factor for polyarthritis in juvenile idiopathic arthritis. Ann Rheum Dis 2008. E-pub ahead of print.

Chapter 4

The *TRAF1/C5* region is a risk factor for polyarthritis in juvenile idiopathic arthritis

F.A.S. Kurreeman
H.M. Albers
J.J. Houwing-Duistermaat
D.M.C. Brinkman
S.S.M. Kamphuis
H.J. Girschick
C. Wouters
M.A.J. Van Rossum
W. Verduijn
R.E.M. Toes
T.W.J. Huizinga
M.W. Schilham
R. ten Cate

Leiden University Medical Center, Department of Paediatrics, Leiden, The Netherlands
Leiden UniversityMedical Center, Department of Rheumatology, Leiden, The Netherlands
Leiden University Medical Center, Department of Medical Statistics, Leiden, The Netherlands
Erasmus Medical Center- Sophia Children's Hospital, Department of Paediatric Immunology/ Rheumatology, Rotterdam, The Netherlands
University of Wuerzburg, Section of Paediatric Rheumatology, Wuerzburg, Germany
University Hospital Gasthuisberg, Department of Paediatric Rheumatology, Leuven, Belgium
Emma Children's Hospital AMC, Department of Paediatrics, Amsterdam, The Netherlands
Leiden University MedicalCenter, Department of Immunohaematology and Bloodbank, Leiden, The Netherlands

Ann Rheum Dis. 67, 1578-1580 (2008).

Abstract

Objective

Juvenile idiopathic arthritis (JIA) is a chronic disorder in which both genetic and environmental factors are involved. Recently, we identified the *TRAF1/C5* region (located on chromosome 9q33–34) as a risk factor for rheumatoid arthritis (RA) ($p_{combined}=1.4x10^{-8}$). In the present study the association of the *TRAF1/C5* region with the susceptibility to JIA was investigated.

Methods

A case-control association study was performed in 338 Caucasian patients with JIA and 511 healthy individuals. We genotyped the single nucleotide polymorphism rs10818488 as a marker for the *TRAF1/C5* region.

Results

The A allele was associated with the susceptibility to rheumatoid factor-negative polyarthritis with an 11% increase in allele frequency (OR 1.54, 95% CI 1.09 to 2.18; p=0.012). This association was stronger when combining subtypes with a polyarticular phenotype (OR 1.46, 95% CI 1.12 to 1.90; p=0.004). In addition, we observed a trend towards an increase in A allele frequency in patients with extended oligoarthritis versus persistent oligoarthritis (49%, 38% respectively); p=0.055.

Conclusions

Apart from being a well replicated riskfactor for RA, *TRAF1/C5* also appears to be a risk factor for the rheumatoid factor-negative polyarthritis subtype of JIA and, more generally, seems to be associated with subtypes of JIA characterised by a polyarticular course.

Introduction

Juvenile idiopathic arthritis (JIA) is defined as arthritis of unknown aetiology that persists for at least 6 weeks and begins before the age of 16 years. It is the most common chronic inflammatory rheumatic disease in childhood (1). In 1997 the International League of Associations for Rheumatology formulated criteria for the classification of seven different subtypes of JIA based on clinical and laboratory features (2, 3).

Although the pathogenesis and aetiology are still poorly understood, JIA is thought to be an autoimmune disease in which genetic and environmental factors play a part. Evidence for the importance of a genetic component includes the ethnic variability in the incidence of different subtypes of JIA, a female preponderance, an increased sibling recurrence rate (λs) of 15 and the association with HLA and non-HLA genes (eg, PTPN22)(3–5). Still little is known about which genetic factors are involved in the susceptibility to JIA and the severity of JIA. Identification of these genetic factors could help to understand the pathogenesis of JIA and could be of use to identify patients with an unfavourable prognosis in an early stage.

Recently we identified the *TRAF1/C5* region (located on chromosome 9q33–34) as a genetic risk factor for RA, using a candidate gene approach (6). A similar finding was made in a genome-wide study in RA (7). *TRAF1* is encoding the *tumour necrosis factor (TNF)-receptor-associated factor 1* and *C5* is encoding the *complement component 5*. Inspired by these results we have studied whether variability in the *TRAF1/C5* region also affects the susceptibility to JIA.

Materials & Methods

A case-control association study was performed in 338 Caucasian patients with JIA from paediatric rheumatology centres in the Netherlands (54%), Belgium (24%) and Germany (22%). Genotype frequencies of 511 healthy unrelated Dutch adult controls were available of the 524 controls previously described by Kurreeman et al due to random genotyping failure (6). All patients with JIA (72% female, 28% male) were diagnosed according to the revised International League of Associations for Rheumatology criteria(2). The inclusion of patients focused on persistent (39%) and extended (14%) oligoarthritis and rheumatoid factor (RF)-negative (22%) and RF-positive (5%) polyarthritis because of their relative homogeneous phenotypes. Patients who were included in the study had a follow-up of at least 2 years. Informed consent from all patients and/or parents and approval from each institutional review board were obtained. DNA was isolated from blood samples (20%) or mouthswabs (80%).

One tagging single nucleotide polymorphism (rs10818488) was genotyped as it revealed the strongest association in RA and none of the other tagging single nucleotide polymorphisms in the 65 kb block provided additional information.6 Rs10818488 is highly linked with rs3761847 (r^2=1, data from HAPMAP) and rs2900180 (r^2=0.66), which were associated with RA as well (7). Rs10818488 was genotyped by the polymerase chain reaction–restriction fragment length polymorphism method as described (6). Each 96-well plate contained two positive and two negative controls. Eight per cent of the samples were run in duplicate and we observed a concordance rate >98%.

Differences in genotype frequencies between cases and controls were assessed using the Cochran–Armitage trend test. Allelic odds ratios (OR) with 95% confidence interval (CI) as well as the genotype specific odds ratios were computed. Case and control genotype frequencies were in Hardy–Weinberg equilibrium. Statistical analysis was performed with SPSS 14.0. p<0.05 was considered statistically significant.

Results

JIA is a heterogeneous disease consisting of several subtypes. As it is best to investigate genetic risk factors in well-defined phenotypic groups, we have analysed the genotypes of the rs10818488 single nucleotide polymorphism in the different subtypes of JIA as well as in the overall group of patients with JIA as shown in Table 1.

Frequencies in patients with persistent oligoarthritis, systemic JIA and in the overall patient group with JIA did not differ significantly fromthose in controls. In extended oligoarthritis and RF-positive polyarthritis (the equivalent of RA) a trend towards an increased A allele frequency was observed (49%, 50% respectively vs 41% in controls). However, although we do observe an 8-9% increase in the A allele, we possess limited power to detect significant differences. In RF-negative polyarthritis patients we found a significant increase in the A allele by 11% (allelic OR 1.54, 95% CI 1.09 to 2.18) when compared with controls. Homozygotes for the susceptibility allele (AA) had an OR of 2.51 (95%CI 1.23 to 5.14) compared with the homozygotes of the protective allele (GG), whereas heterozygotes had an OR of 1.50 (95% CI 0.81 to 2.77). Gender was not a significant covariate when performing a regression analysis (p=0.124).

Table 1 Genotype and allele frequencies in different subtypes of JIA versus controls

		Allele frequency	Genotype frequency				Allelic OR	Genotypic OR* (95% CI)	
Diagnosis†	n	A	AA	AG	GG	p Value‡	(95% CI)	AA	AG
Controls	511	0.41	79 (0.16)	265 (0.52)	167 (0.33)				
Persistent oligoarthritis	133	0.38	18 (0.14)	65 (0.49)	50 (0.38)	0.297	0.87 (0.66 to 1.14)	0.76 (0.42 to 1.39)	0.82 (0.54 to 1.24)
Extended oligoarthritis	48	0.49	10 (0.21)	27 (0.56)	11 (0.23)	0.136	1.36 (0.89 to 2.06)	1.92 (0.78 to 4.71)	1.55 (0.75 to 3.20)
RF-negative polyarthritis	73	0.52	19 (0.27)	38 (0.51)	16 (0.22)	0.012§	1.54 (1.09 to 2.18)	2.51 (1.23 to 5.14)	1.50 (0.81 to 2.77)
RF-positive polyarthritis	18	0.50	5 (0.28)	8 (0.44)	5 (0.28)	0.288	1.42 (0.73 to 2.75)	2.11 (0.60 to 7.51)	1.01 (0.32 to 3.13)
Systemic JIA	41	0.37	3 (0.07)	24 (0.59)	14 (0.34)	0.375	0.82 (0.51 to 1.30)	0.45 (0.13 to 1.62)	1.1 (0.54 to 2.1)
All patients with JIA	338	0.44	59 (0.17)	179 (0.53)	100 (0.30)	0.281	1.11 (0.91 to 1.35)	1.25 (0.82 to 1.90)	1.13 (0.83 to 1.54)

JIA, juvenile idiopathic arthritis; RF, rheumatoid factor.
*GG as reference genotype.
†Diagnosis according to the revision International League of Associations for Rheumatology classification.[2]
‡p Value of the Cochran–Armitage trend test.
§p Value of <0.05 is considered statistically significant.

As extended oligoarthritis, RF-negative polyarthritis and RFpositive polyarthritis have a considerable phenotypic overlap and share a polyarticular course of disease, we also analysed these subtypes grouped together to determine whether the *TRAF1/C5* region predisposes to a polyarticular disease course. The A allele was significantly increased in this combined group (51% in cases, 41% in controls), with an allelic OR 1.46 (95% CI 1.12 to 1.90), a genotypic OR (AA vs GG) of 2.25 (95% CI 1.29–3.90) and a p value (Cochran–Armitage trend test) of 0.004, which remains significant after Bonferroni correction ($p<0.013$).

As persistent and extended oligoarthritis are clinically similar at disease onset, but differ in their course and outcome, we tested the hypothesis that these two subtypes would also differ in their genetic predisposition. Intriguingly, we detected a borderline significant difference between these two subtypes when compared directly with each other (allelic OR 1.57, 95% CI 0.98 to 2.51; $p=0.055$).

Together, these data indicate that the A allele predisposes predominantly to subtypes of JIA characterised by a polyarticular course, indicating that this allele does not associate with JIA as such, but rather with a particular phenotype of JIA.

Discussion

This is one of the first studies to report an association between the *TRAF1/C5* region and JIA. The A allele of rs10818488 was significantly associated with the susceptibility to RF-negative polyarthritis. However, patients with RF-negative polyarthritis have a considerable phenotypic overlap with patients with extended oligoarthritis and RF-positive polyarthritis, in having a polyarticular course of disease. Intriguingly, rs10818488 seems to be associated with this polyarticular phenotype and this difference remained significant after Bonferroni correction for multiple tests ($p<0.013$).

Although we cannot formally exclude the possibility that population stratification may play a role in this study, we did not observe any statistically significant differences in the minor allele frequencies of the patients from the Netherlands, Belgium and Germany. Additionally, analysing the Dutch population independently did not alter the results (eg, RF-negative polyarthritis versus controls: allelic OR 1.66 (95% CI 1.10 to 2.51), genotypic OR (AA versus GG) 3.05 (95% CI 1.25 to 7.44).

At the clinical level, it is also important to make a distinction between persistent oligoarthritis and extended oligoarthritis in order to predict the probability of the development of a polyarticular disease course and adjust the medical treatment accordingly. Comparison of patients with persistent and extended oligoarthritis revealed a borderline significant difference in A allele frequency. This may indicate that the *TRAF1/C5* region may eventually be helpful in predicting the development of an extended course in patients with oligoarthritis. Extended studies of this polymorphism may confirm its relevance as a predictive marker.

It is presently unclear how the association between the *TRAF1/C5* region and disease susceptibility can be explained biologically. The associated polymorphism is highly linked to the TRAF1 gene as well as the 3' untranslated region of the C5 gene. Activated complement component 5 acts as a strong chemoattractant for neutrophils, and a deregulated activity may contribute to the perpetuation of inflammation. In JIA, complement activation is occasionally observed, especially in active polyarthritis (8). On the other hand, TRAF1 plays an essential role in the intracellular TNF-signalling pathway and is possibly a negative regulator of TNF signalling (9). Evidence for the importance of TNF in JIA is suggested by the effectiveness of treatment directed against TNFα, especially in subtypes with a polyarticular course. However, future research will be necessary to confirm the genetic association and investigate functional consequences of the associated allele and could reveal further insight in the pathogenesis of polyarticular disease in JIA. Apart from being a well replicated risk factor for RA, *TRAF1/C5* also appears to be a risk factor for the RF-negative polyarthritis subtype of JIA and, more generally, seems to be associated with subtypes of JIA characterised by a polyarticular course.

References

(1) Manners PJ, Bower C. Worldwide prevalence of juvenile arthritis why does it vary so much? J Rheumatol 2002;29:1520–30.

(2) Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.

(3) Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007;369:767–78.

(4) Saurenmann RK, Rose JB, Tyrrell P, Feldman BM, Laxer RM, Schneider R, et al. Epidemiology of juvenile idiopathic arthritis in a multiethnic cohort: ethnicity as a risk factor. Arthritis Rheum 2007;56:1974–84.

(5) Phelan JD, Thompson SD. Genomic progress in pediatric arthritis: recent work and future goals. Curr Opin Rheumatol 2006;18:482–9.

(6) Kurreeman FAS, Padyukov L, Marques RB, Schrodi SJ, Seddighzadeh M, Stoeken-Rijsbergen G, et al. A candidate gene approach identifies the *TRAF1/C5* region as a risk factor for rheumatoid arthritis. PLoS Med 2007;4:e278.

(7) Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B, et al. TRAF1-C5 as a risk locus for rheumatoid arthritis—a genomewide study. N Engl J Med 2007;357:1199–209.

(8) Jarvis JN. Pathogenesis and mechanisms of inflammation in the childhood rheumatic diseases. Curr Opin Rheumatol 1998;10:459–67.

(9) Tsitsikov EN, Laouini D, Dunn IF, Sannikova TY, Davidson L, Alt FW, et al. TRAF1 is a negative regulator of TNF signaling: enhanced TNF signaling in TRAF1-deficient mice. Immunity 2001;15:647–57.

(10) Lovell DJ, Reiff A, Jones OY, Schneider R, Nocton J, Stein LD, et al. Long-term safety and efficacy of etanercept in children with polyarticular-course juvenile rheumatoid arthritis. Arthritis Rheum 2006;54:1987–94.

Chapter 5

The *TRAF1-C5* region on chromosome 9q33 is associated with multiple autoimmune diseases

F.A.S. Kurreeman
G.N. Goulielmos
B.Z. Alizadeh
B. Rueda
J. Houwing-Duistermaat
E. Sánchez
M. Bevova
T. RRadstake
M.C. Vonk
E.I. Galanakis
N. Ortego
W. Verduyn
M.I. Zervou

SLEGEN Consortium
B.O. Roep
B. Dema
L. Espino
E. Urcelay
D.T. Boumpas
L.H. van den Berg
C. Wijmenga
B.P.C. Koeleman
T.W.J. Huizinga
R.E.M Toes
J. Martin

Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
Department of Medicine, University of Crete, Greece
Department of Biomedical Genetics, Utrecht University Medical Center, Utrecht, The Netherlands
Instituto de Biomedicina López-Neyra, CSIC, Granada, Spain
Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands
Department of Rheumatology, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands
Department of Internal Medicine, Hospital Clínico San Cecilio, Granada, Spain.
Department of Immunohaematology and Bloodbank, Leiden University Medical Center, Leiden, The Netherlands
Department of Immunology, Hospital Clínico San Carlos, Madrid, Spain.
Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands
Department of Human genetics, Groningen University Medical Center and University of Groningen, Groningen, The Netherlands

Ann Rheum Dis. (2009)

Abstract

Objectives

The TRAF1-C5 locus has recently been identified as a genetic risk factor for rheumatoid arthritis. Since genetic risk factors tend to overlap with several autoimmune diseases, we aimed to investigate whether this region is associated with Type I Diabetes (TID), Celiac Disease (CD), Systemic Sclerosis (SSc) and Systemic Lupus Erythematosus (SLE).

Methods

We genotyped the most consistently associated SNP, rs10818488, in a total of 735 T1D, 1049 CD, 367 SSc, 746 SLE and 3494 ethnically and geographically matched healthy individuals. The replication sample set consisted of 99 T1D, 272 SLE patients and 482 healthy individuals from Crete.

Results

We detected significant association of the rs10818488 A allele with T1D (OR 1.14, p=0.027) and SLE (OR 1.16, p=0.016) which was replicated in 99 T1D, 272 SLE patients and 482 controls from Crete (OR 1.64, p=0.002; OR 1.43, p=0.002 respectively). Joint analysis of all T1D (N=961) and all SLE (N=1018) patients compared to 3976 healthy individuals yielded an allelic common OR of 1.19 (p=0.002) and 1.22 ($P=2.6x10^{-4}$) respectively. However, combining our dataset with the T1D sample set from the WTCCC results in a non-significant association (OR 1.06, p=0.087). In contrast, previously unpublished results from the SLEGEN study shows significant association of the same allele (OR 1.19, p=0.0038) with an overall effect of 1.22 ($p=1.02x10^{-6}$) in a total of 1577 SLE patients and 4215 healthy individuals.

Conclusion

We report significant association of the *TRAF1-C5* locus in SLE implying that this region lies in a pathway relevant to multiple autoimmune diseases.

Introduction

The region encompassing Tumour Necrosis Factor (TNF) receptor associated factor 1 (*TRAF1*) and complement component 5 (C5) has recently reported it to be a genetic risk factor involved in rheumatoid arthritis (RA)(1). The robustness of this association is demonstrated by its prevalent risk in Dutch, Swedish, Crete(2) and American populations and is further corroborated by both a genome-wide association study(3) and an extensive fine-mapping study we have undertaken(4). Interestingly one consistent RA association signal defined by rs10818488 or its perfect proxy (rs3761847 and rs7021049, Linkage Disequilibrium R^2>0.98) has been identified in this region by three independent studies and has been further confirmed by a large European family-based study(5). Moreover, an association of the opposite allele of rs3761847 in RA has been reported in the Japanese population but no association has been found in a Korean population(6;7). Additionally, we and others have shown that this polymorphism is associated with juvenile (idiopathic) arthritis(JIA)(8;9), indicating its possible role in other autoimmune diseases.

From a functional perspective, TRAF1 is likely to be a negative regulator of TNF Receptor signaling(10) and C5 is a central component of the complement pathway(11). Both molecules are potent immune mediators and so far the question remains whether this region is restricted to arthritis as such or whether it lies in a biological pathway common to other autoimmune diseases. In the present study we aimed to investigate this hypothesis further by considering the role of this locus in four diseases including Type 1 Diabetes (T1D), Celiac Disease (CD), Systemic Sclerosis (SSc) and Systemic Lupus Erythematosus (SLE).

Methods

Sample sets

DNA was obtained from cohorts of T1D, CD, SSc and SLE from The Netherlands and Spain as well as T1D and SLE patients from Crete. Control cohorts were matched to each specific population of origin. The T1D sample set consisted of 556 white Dutch and 306 Spanish patients. The CD sample set comprised 496 Dutch and 553 Spanish CD patients. The SSc sample set consisted of 138 Dutch and 229 Spanish Systemic Sclerosis patients. The SLE sample set included 161 Dutch and 585 Spanish patients. All patients were Caucasian. A common set of controls were used consisting of 1396 Dutch and 2098 Spanish geographically and ethnically-matched healthy individuals. The replication sample set consisted of 272 SLE, 99 T1D patients and 482 geographically and ethnically matched controls originating from Crete. All patients were diagnosed using the appropriate classification criteria and are described extensively in the Supplementary Methods. Informed consent was obtained from all subjects, and the study was approved by the local ethics committee of each center.

GWAS data

Genotype counts were obtained for rs10118357 (R^2=1 with rs10818488) from the publicly available WTCCC study(12) consisting of 1960 T1D patients and 2930 ethnically matched healthy individuals from the UK population. For the SLE sample set, genotyope counts were obtained from the SLEGEN study(13) for rs3761847 (R^2=1 with rs10818488) from 720 SLE patients and 2337 healthy individuals. These sample sets are described in further detail elsewhere(12;13).

Genotyping

All samples were genotyped using the Taqman assay (Applied Biosystems) according to the manufacturer's instructions. Each plate consisted of at least 8 positive and 8 negative controls. At least 10% of the samples were genotyped in duplo with no discrepancies observed. The Dutch control sets consists of 511 healthy individuals which were previously genotyped(1). An additional 535 Dutch healthy unrelated individuals were genotyped using allele specific kinetic PCR(14) as well as 715 Dutch controls from Utrecht using the Taqman assay. To compare genotyping methods at least 50 samples were genotyped on each platform (Taqman, allele specific kinetic PCR and RFLP) and revealed a concordance rate of >99%.

Statistical analysis

In the controls, the frequencies were in Hardy-Weinberg equilibrium as determined by the observed versus expected genotype counts. Genotype counts were analysed using SPSS version 12.0. Odds ratios and confidence intervals were calculated using Statcalc. Combining odds ratios across sample sets was performed using the Cochran-Mantel-Haenszel test as implemented in EasyMA and the meta package in R (15). No evidence of heterogeneity of risk effect was observed using the Breslow and Day method (P>0.05) when combining OR. All power calculations were performed using Quanto version 1.2 (http://hydra.usc.edu/gxe). P values below 0.05 were considered significant.

Results

T1D, CD, SSc and SLE samples were genotyped from both the Spanish and Dutch populations for the strongest and most consistent association signal in the *TRAF1-C5* region characterized by SNP rs10818488. Since the Dutch study was largely underpowered to detect modest effect sizes (OR~1.2) we opted for a combined analysis which consisted of a total of 3494 controls and 735 T1D, 1049 CD, 367 SSc, 746 SLE patients. This combined dataset enhanced power to ≥85% to detect effect sizes of 1.2 at P<0.05 with the exception of SSc which only achieved 64% power to detect an odds ratio of 1.2.

The frequency of the rs10818488 A allele was significantly increased in T1D patients (OR 1.14; 95% CI 1.02- 1.28; p=0.027, Table 1). Similarly we found a significant difference in the prevalence of the A allele in SLE patients resulting in an OR of 1.16 (95% CI 1.03-1.31, p=0.016). Patients harboring two copies of the A allele had a 1.3 fold (p=0.01) and a 1.4 fold (p=0.04) increased risk for T1D and SLE respectively as compared to those who carried none. We observed no association with CD (Allelic OR 1.07, 95% CI 0.97-1.18, p=0.18) and SSc (Allelic OR 1.02, 95% CI 0.87-1.19, p=0.84). While the possibility remains that the absence of association in SSc may be due to lack of power to detect OR of 1.2 or lower, the direction of association in the individual studies was opposite (Allelic OR, 95% CI; Spanish 1.16, 0.94-1.43; Dutch 0.85, 0.65-1.09), suggesting that the absence of association may be more likely.

We then proceeded to replicate our significant findings in T1D and SLE in the genetically homogeneous population of Crete. Since this largest island of Greece consists of 650,000 inhabitants who share the same genetic and cultural background as well as a common environment, it represents a “geographically isolated” gene pool which may enhance the detection of risk alleles that may be diluted in larger continental populations(16). We observed an 11% increase and an 8% increase in the A allele in T1D and SLE patients, respectively, when compared to controls (Table 2). This resulted in a 1.6 fold increased risk for T1D (p=0.002) and a 1.4 fold increased risk for SLE (p=0.002) in the Crete population. Overall analysis of T1D and SLE in all three datasets (Spanish, Dutch and Greek) reveals a common OR of ~1.2 (p=0.002 and $p=2.6x10^{-4}$ respectively) for both diseases.

We also analysed our data in combination with rs10118357 (perfect proxy of rs10818488, $R^2=1$) for T1D obtained from the Wellcome Trust Case Control Consortium (WTCCC)(12) study as well as a meta-analysis of the effect of SNP rs3761847 (perfect proxy of rs10818488, $R^2=1$) in the SLE Genetics consortium (SLEGEN) study(13) (Table 2). The WTCCC T1D dataset comprising 1960 patients and 2930 healthy individuals displayed no association with an OR of 0.99 (95% CI 0.91-1.08, p=0.837). Combining all four datasets (2794 patients and 6906 controls) generated an overall OR of 1.06 (95% CI 0.99-1.13, p=0.087) indicating that the effect size is very modest and that our study is largely underpowered to detect a significant effect. Previously unpublished data from the SLEGEN study in contrast, show a significant increase in the rs10818488 A allele frequency in 720 SLE patients (43%) as compared to 2337 healthy individuals (39%) (OR 1.19, 95% CI 1.06-1.35, p=0.0038), independently replicating the association of this locus with SLE. Combining all datasets consisting of 1577 SLE patients and 4215 controls shows a significant association with OR 1.22 (95% CI 1.12-1.31, $p=1.02x10^{-6}$).

Table 1. Association of rs10818488 in the *TRAF1-C5* locus with autoimmune diseases.

		Controls	T1D	CD	SSc	SLE
	Initial Sample Set					
Spanish	AA	169	45	84	38	101
	AG	663	150	251	104	276
	GG	564	111	218	87	208
	A(%)	1001(36%)	240(39%)	419(38%)	180(39%)	478(41%)
Dutch	AA	392	106	100	21	28
	AG	1060	305	249	68	84
	GG	646	145	147	49	49
	A(%)	1844(44%)	517(46%)	449(45%)	110(40%)	140(44%)
	OR(95% CI)*	-	1.14 (1.02-1.28)	1.07 (0.97-1.18)	1.02 (0.87-1.19)	1.16 (1.03-1.31)
	P	-	0.027	0.181	0.842	0.016
	Replication Sample Set					
Crete	AA	32	1			26
	AG	214	77			148
	GG	236	21			98
	A(%)	278(29%)	79(40%)			200(37%)
	OR (95% CI)		1.64 (1.18-2.28)			1.43 (1.14-1.80)
	P		0.002			0.002
	Joint Sample Set					
	Combined OR¥ (95% CI)		1.19 (1.07-1.32)			1.22 (1.09-1.35)
	P		0.002			$2.6x10^{-4}$

T1D – Type 1 Diabetes, CD- Celiac Disease, SSc- Systemic Sclerosis and SLE- Systemic Lupus Erythematosus. * Represents the combined allelic OR between the Dutch and Spanish sample set. ¥ Represents the combined allelic OR between Dutch, Spanish and Crete sample sets. All OR were combined using the Mantel-Haenszel fixed effects as implemented in EasyMA and the Meta package in R.

Table 2. Meta-analysis of rs10818488 in the *TRAF1-C5* locus with Genome Wide Association Studies.

	WTCCC		SLEGEN	
GWAS Data Set	Controls	T1D	Controls	SLE
AA	362	598	339	138
AG	1011	1410	1135	344
GG	587	922	863	238
A (%)	2930 (44%)	1960 (44%)	2337 (39%)	720 (43%)
OR * (95% CI)		0.99 (0.91-1.08)		1.19 (1.06-1.35)
P		0.837		0.0038
Joint Sample Set				
Combined OR¥ (95% CI)		1.06 (0.99-1.13)		1.22 (1.12-1.31)
P		0.087		1.02×10^{-6}

T1D – Type 1 Diabetes and SLE- Systemic Lupus Erythematosus. * Represents the combined allelic OR between the cases and controls in the GWAS sample set. ¥ Represents the combined allelic OR between Dutch, Spanish, Crete and GWAS sample sets. All OR were combined using the Mantel-Haenszel fixed effects as implemented in EasyMA and the Meta package in R.

Discussion

We report here for the first time reproducible association of the *TRAF1-C5* region with SLE complementing the already consistent finding of this variant with RA and juvenile arthritis. Remarkably, the same allele that predisposes to RA and juvenile arthritis also predisposes to SLE lending support to the hypothesis that this region may contribute to a shared pathway involved in RA, JIA and SLE. While the possibility also remains that other additional alleles at this locus may be involved in these diseases, complementary studies undertaking further fine-mapping as well as sequencing will yield further insight into the most likely causal alleles at this locus.

We also observed a difference in allele frequencies in healthy populations of Dutch (44%), Spanish (36%) and Greek (29%) origins. To address this difference in population, each patient-control sample set was geographically and ethnically matched. Data from HapMap (www.hapmap.org) support these observations with the G allele (minor allele in Caucasians) frequency of rs3761847 (perfect proxy of rs10818488) varying from 48% in Caucasians of European descent, 31% in Gujarati Indians in Texas, 42% in Japanese in Tokyo, 66% in Mexicans in Los Angeles, 59% in Yorubans in Nigeria and 74% in Luhya in Kenya. However, underlying population stratification is difficult to account for as panels of markers characterizing each population in our study thoroughly have not as yet been described. In our study, association of the *TRAF1-C5* locus has been observed in more than one sample set considerably reducing the chances of false positive findings due to population stratification.

The association of the *TRAF1-C5* region has not been reported in the recently published genome-wide association studies (GWAS) in SLE(13;17;18) and the GWAS of T1D in WTCCC study(19). However, while an association with RA was initially not detected in the WTCCC , a follow-up report showed a significant association of the same allele in the UK population although with much a much lower effect size(20). In line with this observation, a meta-analysis of GWA studies in RA has recently illustrated that much larger sample sizes are required to detect the commonly observed low penetrance of genes in autoimmune diseases like RA(21). It is also not uncommon that modest effects are not detected in such largescale studies as exemplified by the association of TNFSF4, a gene identified by family-based association studies(22), that did not surface in the recently performed GWAS. Likewise BANK1, a gene identified in the GWA scan by Kozyrev et al was not identified by the other two GWA studies that employed 500K SNPs. However, while our study combined with the GWA of the WTCCC does not provide sufficient evidence of the role of the *TRAF1-C5* locus in T1D, data obtained from the SLEGEN study confirms our findings of this locus in SLE. Interestingly, this association is absent in a well-powered Japanese case-control study and small study in the Columbian population(6;23). As with many genetic loci, it is highly likely that ethnic differences exist in the contribution of this locus to SLE.

TRAF1 has been reported to be a negative regulator of the TNF receptor signalling cascade(10;24) and high levels of TNF has been detected in both human and murine lupus, it is likely that a dysregulation of the function and/or expression of this molecule could be involved in the inflammatory processes in SLE(10;25). Currently, controlled clinical trials involving the blockade of TNF are being conducted(26). Although the haplotype block encompassing rs10818488 does not encompass the C5 coding region(4), this central component of the complement system is also a likely functional candidate in SLE. The prominent role of complement activation in humans and murine models of SLE as well as the beneficial effect of blocking C5 anaphylatoxin in murine lupus models point to the likely role of this molecule in SLE(27;28). In summary, we report here for the first time an association of the *TRAF1-C5* locus with SLE which, in combination with previous findings of an increased risk to RA and juvenile arthritis, indicates that this region is likely to be part of a shared mechanism underlying several autoimmune diseases.

References

(1) Kurreeman FA, Padyukov L, Marques RB, Schrodi SJ, Seddighzadeh M, Stoeken-Rijsbergen G et al. A candidate gene approach identifies the *TRAF1*/C5 region as a risk factor for rheumatoid arthritis. PLoS Med 2007; 4(9):e278.

(2) Zervou MI, Sidiropoulos P, Petraki E, Vazgiourakis V, Krasoudaki E, Raptopoulou A et al. Association of a *TRAF1* and a STAT4 gene polymorphism with increased risk for rheumatoid arthritis in a genetically homogeneous population. Hum Immunol 2008; 69(9):567-71.

(3) Plenge RM, Seielstad M, Padyukov L, Lee AT, Remmers EF, Ding B et al. *TRAF1-C5* as a risk locus for rheumatoid arthritis--a genomewide study. N Engl J Med 2007; 357(12):1199-209.

(4) Chang M, Rowland CM, Garcia VE, Schrodi SJ, Catanese JJ, van der Helm-van Mil AH et al. A large-scale rheumatoid arthritis genetic study identifies association at chromosome 9q33.2. PLoS Genet 2008; 4(6):e1000107.

(5) Kurreeman FA, Rocha D, Houwing-Duistermaat J, Vrijmoet S, Teixeira VH, Migliorini P et al. Replication of the tumor necrosis factor receptor-associated factor 1/complement component 5 region as a susceptibility locus for rheumatoid arthritis in a European family-based study. Arthritis Rheum 2008; 58(9):2670-4.

(6) Nishimoto K, Kochi Y, Ikari K, Yamamoto K, Suzuki A, Shimane K et al. Association study of *TRAF1-C5* polymorphisms with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in Japanese. Ann Rheum Dis 2009.

(7) Lee HS, Korman BD, Le JM, Kastner DL, Remmers EF, Gregersen PK et al. Genetic risk factors for rheumatoid arthritis differ in Caucasian and Korean populations. Arthritis Rheum 2009; 60(2):364-71.

(8) Behrens EM, Finkel TH, Bradfield JP, Kim CE, Linton L, Casalunovo T et al. Association of the *TRAF1-C5* locus on chromosome 9 with juvenile idiopathic arthritis. Arthritis Rheum 2008; 58(7):2206-7.

(9) Albers HM, Kurreeman FA, Houwing-Duistermaat JJ, Brinkman DM, Kamphuis SS, Girschick HJ et al. The *TRAF1*/C5 region is a risk factor for polyarthritis in Juvenile Idiopathic Arthritis. Ann Rheum Dis 2008.

(10) Wajant H, Henkler F, Scheurich P. The TNF-receptor-associated factor family: scaffold molecules for cytokine receptors, kinases and their regulators. Cell Signal 2001; 13(6):389-400.

(11) Schmidt RE, Gessner JE. Fc receptors and their interaction with complement in autoimmunity. Immunol Lett 2005; 100(1):56-67.

(12) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447(7145):661-78.

(13) Harley JB, Alarcon-Riquelme ME, Criswell LA, Jacob CO, Kimberly RP, Moser KL et al. Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXK, KIAA1542 and other loci. Nat Genet 2008; 40(2):204-10.

(14) Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004; 75(2):330-7.

(15) Cucherat M, Boissel JP, Leizorovicz A, Haugh MC. EasyMA: a program for the meta-analysis of clinical trials. Comput Methods Programs Biomed 1997; 53(3):187-90.

(16) Arnaiz-Villena A, Iliakis P, Gonzalez-Hevilla M, Longas J, Gomez-Casado E, Sfyridaki K et al. The origin of Cretan populations as determined by characterization of HLA alleles. Tissue Antigens 1999; 53(3):213-26.

(17) Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, Garnier S et al. Association of Systemic Lupus Erythematosus with C8orf13-BLK and ITGAMITGAX. N Engl J Med 2008.

(18) Kozyrev SV, Abelson AK, Wojcik J, Zaghlool A, Linga Reddy MV, Sanchez E et al. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. Nat Genet 2008; 40(2):211-6.

(19) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447(7145):661-78.

(20) Barton A, Thomson W, Ke X, Eyre S, Hinks A, Bowes J et al. Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility. Hum Mol Genet 2008.

(21) Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 2008; 40(10):1216-23.

(22) Graham DS, Graham RR, Manku H, Wong AK, Whittaker JC, Gaffney PM et al. Polymorphism at the TNF superfamily gene TNFSF4 confers susceptibility to systemic lupus erythematosus. Nat Genet 2008; 40(1):83-9.

(23) Palomino-Morales RJ, Rojas-Villarraga A, Gonzalez CI, Ramirez G, Anaya JM, Martin J. STAT4 but not *TRAF1*/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians. Genes Immun 2008; 9(4):379-82.

(24) Tsitsikov EN, Laouini D, Dunn IF, Sannikova TY, Davidson L, Alt FW et al. *TRAF1* is a negative regulator of TNF signaling. enhanced TNF signaling in *TRAF1*-deficient mice. Immunity 2001; 15(4):647-57.

(25) Aringer M, Smolen JS. The role of tumor necrosis factor-alpha in systemic lupus erythematosus. Arthritis Res Ther 2008; 10(1):202.

(26) Aringer M, Smolen JS. TNF inhibition in SLE: where do we stand? Lupus 2009; 18(1):5-8.

(27) Manderson AP, Botto M, Walport MJ. The role of complement in the development of systemic lupus erythematosus. Annu Rev Immunol 2004; 22:431-56.

(28) Walport MJ. Complement and systemic lupus erythematosus. Arthritis Res 2002; 4 Suppl 3:S279-S293.

Part II

The role of interleukin 10 (IL10) genetic variants in regulating expression levels and disease pathogenesis

Chapter 6

The Role of Interleukin 10 Promoter Polymorphisms in the Susceptibility of Distal Interphalangeal Osteoarthritis

N. Riyazi

F.A.S. Kurreeman

T.W.J. Huizinga

F.W. Dekker

G. Stoeken-Rijsbergen

M. Kloppenburg

Department of Rheumatology
Department of Clinical Epidemiology
Leiden University Medical Center, Leiden, The Netherlands.

J Rheumatol. 32, 1571-1575 (2005)

Abstract

Objective

The interleukin (IL)-10 single nucleotide promoter polymorphism (SNP) –2849A is associated with decreased IL-10 production as measured by lipopolysaccharide (LPS) stimulated whole blood cultures. A low innate production of IL-10 using the same assay is associated with an increased risk of familial osteoarthritis (OA). We investigated the association of 7 novel SNP located downstream of the IL-10 transcription start site: –2849,–2763, –1330, –1082, –819, and –592, constituting the 4 ancient haplotypes, with distal interphalangeal (DIP) OA.

Methods

The study population comprised consecutive patients with and without radiological DIP OA (Kellgren-Lawrence score of ≥2 in one joint) aged 40–70 years from a cohort of subjects with different types of arthritis in an early stage referred to an Early Arthritis Clinic (EAC). DNA typing for IL-10 SNP as well as radiographs of the hands were performed at clinic enrolment. Patients with rheumatoid arthritis, systemic lupus erythematosus, spondyloarthropathies, and psoriatic arthritis were excluded.

Results

The distribution of DIP OA and IL-10 SNP were comparable to representative samples of the Dutch population. In the cohort of 172 subjects, 57 had DIP OA (33%) and 115 (67%) had no DIP OA. No significant association was found between DIP OA and IL-10 SNP and the 4 common haplotypes IL10.1, IL10.2, IL10.3, and IL10.4.

Conclusion

Our data suggest that IL-10 SNP, including –2849, which is associated with differential production, do not play a major role in the susceptibility of DIP OA.

Introduction

Distal interphalangeal (DIP) osteoarthritis (OA) is one of the most frequent subtypes of OA, causing pain and loss of function in the hands. Familial aggregation and twin studies have shown DIP OA to have a strong familial component(1–3). Despite studies focusing on the genetics of DIP OA, the genes involved have not been identified. Several chromosomal regions are reported to be associated with DIP OA, namely quantitative trait loci (QTL) on chromosomes 2q, 7p, and 11q implicated by linkage studies (4), HLA-DR2 by candidate gene analysis in several populations (5,6), and further, a locus on chromosome 2 containing the matrilin-3 gene, by a genomewide scan in an Icelandic population (7).

Loss of articular cartilage and changes in the subchondral bone and joint margins is a hallmark of OA. An increased matrix catabolism characterized by an upregulation of metalloproteinases (MMP) and the depletion of structural macromolecules such as proteoglycans contributes to the OA disease process (8). Increasing data support the role of cytokines in these processes. Although proinflammatory cytokines have been shown to play a pivotal role in the initiation and development of OA, a shift in the balance between the pro and antiinflammatory cytokines is believed to contribute to the loss of integrity of articular cartilage (9).

Among the antiinflammatory cytokines, interleukin (IL)10 appears to be a crucial factor in inflammatory processes(10). In a mouse model of arthritis, IL-10 administration prevents cartilage destruction by reducing IL-1ß and tumor necrosis factor (TNF)-α mRNA expression in articular chondrocytes, while neutralizing anti-IL-10 antibodies accelerate the onset and enhance the severity of arthritis(11).

The human IL-10 gene is highly polymorphic. Gibson, et al (12) identified 7 novel single nucleotide polymorphisms (SNP) in the distal region of the IL-10 promoter and found that certain haplotypes are significantly associated with high or low IL-10 production. Studies have shown that there are striking differences between healthy individuals in their ability to produce IL-10 following lipopolysaccharide (LPS) stimulation of whole blood cultures ex vivo. Moreover it hasbeen shown that IL-10 haplotypes dictate IL-10 production as measured by this assay (13). Of the IL-10 promoter polymorphisms, –2849A, encoded on haplotype IL-10.01, has been shown to correlate best with protein production (13). Individuals with the AA genotype have a reduced IL-10 production in comparison to AG and GG individuals.

Several studies have reported the clinical relevance of IL10 gene polymorphism in autoimmune diseases such as systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA)(12,14), and in infectious diseases (15). In an earlier study we found an increased risk of familial OA at multiple sites in subjects with a low innate production of IL-10, as measured by the same ex vivo whole blood assay using LPS stimulation(16) used in all previous studies (12,13). This finding implicating IL-10 in cartilage destruction led us to investigate whether genetic variation in IL-10 contributes to the susceptibility of DIP OA. Our aim was to investigate whether IL-10 SNP –592, –819, –1082, –2763, –2849, and –3575 that constitute the 4 ancient IL-10 haplotypes are a risk factor for DIP OA.

Materials & Methods

Study population

The study population consists of subjects from the outpatient clinic of the Department of Rheumatology between the ages of 40 to 70 years with and without radiological DIP OA. This population, enrolled between 1993 and 2000, is part of an ongoing project, the Early Arthritis Clinic (EAC). Consecutive patients with arthritis in at least one joint with a short history of complaints are admitted to this clinic by general practitioners. Each patient subsequently undergoes full clinical, biochemical, and radiographic assessment. Diagnoses in the EAC are made according to international classification and if necessary revised up to 1 year of followup (17). Patients with a definitive EAC diagnosis after one year of followup were included in this study. Patients with RA, SLE, and spondyloarthropathies were excluded because these diseases have been associated with IL-10. Patients with psoriatic arthritis were also excluded since DIP involvement is common in psoriatic arthritis and could thus interfere with our readings.

Reference populations

Prevalence of radiological DIP OA and distribution of IL-10 promoter polymorphisms in the study population were compared to 2 reference populations representative of the general Dutch population. As a reference for radiological DIP OA, data were used from the Zoetermeer population, a population survey consisting of 3109 men and 3476 women (18) as described (5). A comparison between the frequency of radiological DIP OA in the present population in reference to the Zoetermeer population gave an observed-to-expected ratio of 0.96 (0.7–1.2). Distribution of IL-10 promoter polymorphisms was compared to its distribution in a random panel of the southwest region of the Netherlands (n = 321)(19).

Radiographs and radiographic scoring

Plain dorsovolar hand radiographs were taken routinely in each patient during the period of the first visit to the EAC. Only radiographs obtained at a maximum of 3 months before till 3 months after the first visit were included. Each of the radiographs was independently graded using the Kellgren-Lawrence scale for DIP OA by 2 out of 3 observers blinded to the underlying EAC diagnosis. This overall score distinguishes 5 degrees of severity of OA according to the presence of radiological features: osteophytes, joint space narrowing (JSN), subchondral sclerosis, cysts, and deformity. A patient was diagnosed with DIP OA if a Kellgren score of 2 or more was observed in at least one DIP joint. The inter-rater agreement for the presence or absence of DIP OA was 0.7 (Cohen's kappa). In case of disagreement radiographs were re-evaluated until consensus was reached.

Determination of IL-10 promoter polymorphisms

Peripheral blood cells were lysed using sodium dodecyl sulfate and treated with proteinase K. DNA was isolated by phenol-chloroform extraction. A Perkin-Elmer thermal cycler Gene Amp 9600 (PE-Cetus, Norwalk, CA, USA) was used to amplify the IL-10 promoter region by polymerase chain reaction. The primer combination and methods used have been described (20).

Statistical analysis

Means were compared using an independent sample Student's t test. Odds ratios (OR) with 95% confidence intervals (95% CI) were calculated to determine whether the distribution of IL-10 promoter polymorphisms was comparable to the random panel by comparing the distribution of the minor allele grouped with the heterozygote [11 + 12] versus the major allele [22] in both sets of populations. OR were also used to assess the association between DIP OA and the different genotypes of IL10 promoter polymorphisms. Multiple locus haplotype frequencies (SNPHAP: a program for estimating frequencies of haplotypes of large numbers of diallelic markers from unphased genotype data from unrelated subjects (21) and the measures of pairwise linkage disequilibrium (LD) were determined using the HAPLO program (22). It uses an expectation-maximization algorithm to calculate maximum likelihood estimates of haplotype frequencies, given genotype measurements. The estimator of LD (where $D = h_{pq} - pq$) indicates the difference in the observed (h_{pq}) and expected (pq) frequencies of haplotypes. Its maximum value depends on the allele frequencies and whether the rare alleles are associated together on a haplotype (positive value of D) or whether the common allele is associated with the rare allele (negative value of D). |D| (Lewontins D') is the fraction of D of its maximum ($D_{max} = p - pq$) or minimum ($D_{min} = -pq$) possible value. Power analysis was based on achieving 5% significance in order to detect a difference if IL-10 promoter polymorphisms or IL-10 haplotypes were associated with an almost 3-fold increased risk of DIP OA (OR = 2.7). This analysis was based on the frequency of the IL-10 A carriage rate in the SNP –2849 associated with differential IL-10 production.

Results

Six hundred and one consecutive patients aged between 40 to 70 years visited the EAC. Sixty-nine patients were excluded because no definitive diagnosis was made within one year, 281 patients were excluded based on diagnosis of RA, systemic diseases, spondyloarthropathies, and psoriatic arthritis and 79 patients were excluded due to missing data. Of the remaining 172 patients included in the study, 3 were not genotyped for the –3575, –2763, –1082, and –819 SNP due to poor DNA quality and an additional 18 samples were missing at the time of genotyping the –3575 SNP.

Patient characteristics

Because the patients in this study were included in the EAC, they had a broad variety of diagnoses independent of the presence or absence of DIP OA. Fifty-seven (33%) patients had DIP OA. As expected, the average age of patients with DIP OA was higher than patients with no DIP OA [58 yrs (range 43–70) vs 51 yrs (40–70); mean difference: 6.8 (4.1–9.1)]. The clinical characteristics of the study population are summarized in Table 1.

Distribution of IL-10 promoter polymorphisms in comparison to controls

In Table 2, the distribution of the minor allele grouped with the heterozygote is shown versus the major allele of the IL-10 promoter polymorphisms –3575, –2849, –2763, –1082, and –819 in the study population and in the random panel. No difference was observed in the distribution of these alleles in the 2 groups. The IL-10 promoter polymorphisms –1330 and –592 are in complete linkage disequilibrium with –1082 and –819 respectively; therefore, these variables were excluded from further analyses of single promoter polymorphisms. Genotypes did not show deviations from the Hardy-Weinberg equilibrium (data not shown) (23).

Association of IL-10 promoter polymorphisms and haplotypes with DIP OA

The association between DIP OA and genotypes of IL-10 SNP –3575, –2849, –2763, –1082, and –819 is shown in Table 3. No association was found between these polymorphisms and DIP OA. Since the interaction of 2 or more SNP in haplotypes may be more informative than single polymorphisms, a haplotype analysis was performed. No difference was observed in the distribution of the 4 extended haplotypes, IL-10.1, IL10.2, IL-10.3, and IL-10.4 in patients with and without DIP OA ($p = 0.67$). Data are presented in Table 4. In a separate analysis of the distal haplotype frequency (A–3575T, A–2849G, and A–2763C) and the proximal haplotype frequency (A-1330G, G-1082 A, T-819 C, and A-592 C) no difference was seen in the distribution in patients with and without DIP OA (data not shown).

Table 1. Clinical characteristics of the study population: patients with and without distal interphalangeal (DIP) osteoarthritis (OA).

	DIP OA (n = 57)	No DIP OA (n = 115)
Mean age (range)	58 (43–70)	51 (40–70)
Women, n (%)	33 (58)	55 (47)
EAC diagnoses, n		
Septic arthritis	1	2
Reactive arthritis	3	7
Crystal arthropathy	8	18
Post-traumatic	1	4
Osteoarthritis	17	10
Unclassified arthritis	23	51
Malignancy related arthritis	1	5
Sarcoidosis	1	5
Other	2	11
Unknown	0	2

Table 2. Distribution of the minor allele grouped with the heterozygote (11 + 12) versus the major allele (22) in single nucleotide promoter polymorphisms (SNP) of the interleukin (IL)-10 gene in the study population compared to a random panel in the south-west region of the Netherlands expressed as odds ratios (OR) with 95% confidence intervals (95% CI).

SNP	11 + 12	22	OR (95% CI)
IL-10 A-3575T			
Study population	0.71	0.29	
Random panel	0.64	0.36	1.4 (0.7–2.7)
IL-10 A-2849G			
Study population	0.55	0.45	
Random panel	0.43	0.57	1.6 (0.9–3.0)
IL-10 A-2763C			
Study population	0.65	0.35	
Random panel	0.58	0.42	1.3 (0.7–2.5)
IL-10 G-1082A			
Study population	0.75	0.25	
Random panel	0.76	0.24	1.0 (0.5–1.9)
IL-10 T-819C			
Study population	0.40	0.60	
Random panel	0.40	0.60	1.0 (0.6–1.8)

Table 3. Summary of the association between DIP OA (n = 57) and the genotypes of the IL-10 promoter polymorphisms within the study population (n = 172) expressed as OR (95% CI).

SNP	Genotype distribution			11 versus 12 + 22	11 + 12 versus 22
	11	12	22	OR (95% CI)	OR (95% CI)
IL-10 A-3575T*					
DIP OA	9	27	14		
No DIP OA	13	57	30	1.5 (0.5–4.1)	1.1 (0.5–2.5)
IL-10 A-2849G					
DIP OA	4	30	23		
No DIP OA	8	52	55	1.0 (0.2–3.9)	1.4 (0.7–2.7)
IL-10 A-2763C*					
DIP OA	7	33	16		
No DIP OA	12	58	43	1.3 (0.4–3.8)	1.5 (0.7–3.3)
IL-10 G-1082A*					
DIP OA	16	26	15		
No DIP OA	22	62	28	1.6 (0.7–3.6)	0.93 (0.4–2.1)
IL-10 T-819C*					
DIP OA	3	18	36		
No DIP OA	6	41	65	0.98 (0.2–4.7)	0.81 (0.4–1.6)

Table 4. Haplotype frequencies in patients with DIP OA (n = 57) and patients without DIP OA (n = 116).

IL10 HAP	-3575	-2849	-2763	-1330	-1082	-819	-592	DIP OA	No DIP OA
IL10.1	A	A	A	A	G	C	C	0.30	0.23
IL10.2	T	G	C	G	A	C	C	0.28	0.27
IL10.3	A	G	A	A	G	C	C	0.19	0.23
IL10.4	T	G	C	G	A	T	A	0.07	0.09

Discussion

We are the first to investigate the relationship between a highly genetic form of OA, namely DIP OA and promoter polymorphisms located downstream of the IL-10 transcription start site: –2849, –2763, –1330, –1082, –819, and –592 constituting the 4 ancient IL-10 haplotypes. These SNP as well as the 4 haplotypes were not associated with a higher risk of radiological DIP OA. Association of DIP OA with IL-10 promoter polymorphisms in our investigation was studied in patients with a variety of underlying forms of arthritis. The consecutive patient population was collected in a prospective manner. Because the study population consisted of patients included in an EAC, the existing correlation between certain rheumatic diseases and the genetic variables under study was taken into consideration while selecting patients. All patients with diseases reported in the literature to be associated with the polymorphisms under study were excluded. Furthermore, patients with diseases that can lead to radiological damage of the DIP joints, such as psoriatic arthritis, were also excluded. However, radiological damage would not have been very likely since patients are seen in the EAC at a very early stage. Although patient selection in the EAC was outside our research question, we did find a patient population with a well-documented phenotype, i.e., radiological DIP OA with IL-10 and DIP OA distributions comparable to the general population.

Among the IL-10 promoter polymorphisms under study, the SNP –2849A has been associated with low IL-10 production. We found an OR of one when comparing the –2849 AA versus –2849AG and –2849GG. These data strongly suggest that the association is absent; although a type II error may have occurred given the power of the current study. We conclude that it is unlikely that the SNP –2849A has a large effect on genetic susceptibility to radiological DIP OA.

Although the antiinflammatory role of IL-10 is recognized in arthritis, its role in OA is still undefined. Based on its biologic activity it is, however, conceivable that low IL10 production may contribute to a catabolic state in OA. IL10 is an important immunoregulatory cytokine in man (10) and plays a crucial role in inflammation and tissue destruction. In an arthritis model, mice lacking the gene for IL-10 experienced higher rates of clinical signs and more severe knee and paw injury compared to IL-10 wild-type controls. Furthermore, plasma levels of TNF-a, IL-1ß, and IL-6 were also enhanced in knockout compared to wild-type mice (24).

IL-10 has been reported in vitro to be synthesized in increased amounts either spontaneously by synovial membrane and cartilage (25) or after stimulation of chondrocytes with IL-1ß or TNF-α (8). IL-10 contributes to cartilage homeostasis through several pathways. In experiments on joint tissue, it has been shown that a lack of IL-10 can lead to joint destruction as a result of an increased expression of metalloproteinases (26). Upregulation of IL-1 receptor antagonist production by isolated monocytes has been found for IL-10 by human monocytes and neutrophils (27). Besides exerting antiinflammatory activity, IL-10 has been shown to directly stimulate proteoglycan synthesis by human chondrocytes in vitro (28).

In an earlier study (16), we investigated whether genetic variation in cytokine production contributes to OA susceptibility. The ex vivo production of cytokines in whole blood assays was used to classify subjects as “high” or “low” producers based on the distribution in controls. We observed that a low innate production of IL-10 using LPS production was associated with an increased risk of familial OA at multiple sites (defined as multiple sites in the hands or at 2 or more joint sites including the hands, spine, knees, and hips). The low innate IL-10 production in these patients is assumed to be partly caused by genetic variation at the IL10 locus (13,15), thus implicating that genetic variation of ex-vivo cytokine production upon LPS stimulation contributes to OA susceptibility. Therefore, we investigated whether IL10 SNP that constitute the 4 ancient IL-10 haplotypes, some associated with IL-10 production, are a risk factor for a predominant subtype of OA specifically in the DIP joints. Our results indicate that this is not the case. The discrepancy between findings in our earlier study and the present study may be due to different OA phenotypes: familial OA at multiple joint sites versus radiographically defined OA in DIP joints. Alternatively, not all of the genetic variation in IL-10 production is dictated by IL-10 haplotypes (13). It may be that genetic factors that dictate inter-individual IL-10 production differences that are not located in the IL-10 locus are those relevant for DIP OA.

In summary our current work suggests that the currently known IL-10 SNP do not exhibit a major effect on genetic susceptibility to DIP OA.

References

(1) Stecher RM. Heberden's nodes: heredity in hypertrophic arthritis of finger joints. Am J Med Sci 1941;201:801-9.

(2) Spector TD, Ciccuttini F, Baker J, Loughlin J, Hart DJ. Genetic influences on osteoarthritis in women: a twin study. BMJ 1996;312:940-3.

(3) Jonsson H, Manolescu I, Stefansson SE, et al. The inheritance of hand osteoarthritis in Iceland. Arthritis Rheum 2003;48:391-5.

(4) Spector TD, MacGregor AJ. Risk factors for osteoarthritis: genetics. Osteoarthritis Cartilage 2004;12:S39-44.

(5) Riyazi N, Spee J, Huizinga TWJ, et al. HLA class II is associated with distal interphalangeal osteoarthritis. Ann Rheum Dis 2003;62:227-30.

(6) Merlotti D, Santacroce C, Gennari L, et al. HLA antigens and primary osteoarthritis of the hand. J Rheumatol 2003;30:1298-304.

(7) Stefansson SE, Jonsson H, Ingvarsson T, et al. Genomewide scan for hand osteoarthritis: a novel mutation in matrilin-3. Am J Hum Genet 2003;72:1448-59.

(8) Hedbom E, Häuselmann HJ. Molecular aspects of pathogenesis in osteoarthritis: the role of inflammation. Cell Mol Life Sci 2002;59:45-53.

(9) Martel-Pelletier J, Alaaeddine N, Pelletier JP. Cytokines and their role in the pathophysiology of osteoarthritis. Front Biosci 1999;4:694-703.

(10) Fickenscher H, Hor S, Kupers H, Knappe A, Wittmann S, Sticht H. The interleukin-10 family of cytokines. Trends Immunol 2002;23:89-96.

(11) Joosten LA, Lubberts E, Durez P, et al. Role of interleukin-4 and interleukin-10 in murine collagen-induced arthritis. Protective effect of interleukin-4 and interleukin-10 treatment on cartilage destruction. Arthritis Rheum 1997;40:249-60.

(12) Gibson AW, Edberg JC, Wu J, Westendorp RG, Huizinga TW, Kimberly RP. Novel single nucleotide polymorphisms in the distal IL-10 promoter affects IL-10 production and enhance the risk of systemic lupus erythematosus. J Immunol 2001;166:3915-22.

(13) Kurreeman FA, Schonkeren JJ, Heijmans BT, Toes RE, Huizinga TW. Transcription of the IL10 gene reveals allele specific regulation at the mRNA level. Hum Mol Genet 2004;13:1755-62.

(14) Huizinga TW, Keijsers V, Yanni G, et al. Are differences in interleukin 10 production associated with joint damage? Rheumatology Oxford 2000;39:1180-8.

(15) Westendorp RG, Langermans JAM, Huizinga TW, et al. Genetic influence on cytokine production and fatal meningococcal disease. Lancet 1997;349:170-3.

(16) Riyazi N, Slagboom E, de Craen AJM, et al. Association of the risk of osteoarthritis with high innate production of interleukin–1beta and low innate production of interleukin-10 ex vivo, upon lipopolysaccharide stimulation. Arthritis Rheum 2005;52:1443-50.

(17) Van der Horst-Bruinsma IE, Speyer I, Visser H, Breedveld FC, Hazes JMW. Diagnosis and course of early onset arthritis: results of a special early arthritis clinic compared to routine patient care. Br J Rheumatol 1998;37:1084-8.

(18) Van Saase JLCM, Vandenbroucke JP, van Romunde LKJ, Valkenburg HA. Osteoarthritis and obesity in the general population: a relationship calling for an explanation. J Rheumatol 1988;15:1152-8.

(19) de Jong BA, Westendorp RG, Eskdale J, Uitdehaag BM, Huizinga TW. Frequency of functional interleukin-10 promoter polymorphism is different between relapse-onset and primary progressive multiple sclerosis. Hum Immunol 2002;63:281-5.

(20) Lard LR, van Gaalen FA, Schonkeren JJ, et al. Association of the -2849 interleukin-10 promoter polymorphism with autoantibody production and joint destruction in rheumatoid arthritis. Arthritis Rheum 2003;48:1841-8.

(21) Clayton, D. Haplotyping and TDT programs [homepage on the Internet]. MRC Biostatistics Unit, Institute of Public Health, Cambridge Institute for Medical Research, Cambridge University, UK; Nov 2000 [accessed 2005 April 26]. Available from: http://www-gene.cimr.cam.ac.uk/clayton/software/

(22) Long JC, Williams RC, Urbanek M. An E-M algorithm and testing strategy for multiple-locus haplotypes. Am J Hum Genet 1995;56:799-810.

(23) Nielsen DM, Ehm MG, Weir BS. Detecting marker-disease association by testing for Hardy-Weinberg disequilibrium at a marker locus. Am J Hum Genet 1998;63:1531-40.

(24) Cuzzocrea S, Mazzon E, Dugo L, et al. Absence of endogenous interleukin-10 enhances the evolution of murine type-II collagen-induced arthritis. Eur Cytokine Netw 2001;12:568-80.

(25) Pelletier JP, Martel-Pelletier J, Abramson SB. Osteoarthritis, an inflammatory disease: potential implication for the selection of new therapeutic targets. Arthritis Rheum 2001;44:1237-47.

(26) van Roon JA, Lafeber FP, Bijlsma JW. Synergistic activity of interleukin-4 and interleukin-10 in suppression of inflammation and joint destruction in rheumatoid arthritis. Arthritis Rheum 2001; 44:3-12.

(27) Jenkins JK, Malyak M, Arend WP. The effects of interleukin-10 on interleukin-1 receptor antagonist and interleukin-1 beta production in human monocytes and neutrophils. Lymphokine Cytokine Res 1994;13:47-54.

(28) van Roon JA, van Roy JL, Gmelig-Meyling FH, Lafeber FP, Bijlsma JW. Prevention and reversal of cartilage degradation in rheumatoid arthritis by interleukin-10 and interleukin-4. Arthritis Rheum 1996;39:829-35.

Chapter 7

Transcription of the *IL10* gene reveals allele-specific regulation at the mRNA level

F.A.S. Kurreeman

J.J.M. Schonkeren

B.T. Heijmans

R.E.M. Toes

T.W.J. Huizinga

Department of Rheumatology, Leiden University Medical Centre, PO Box 9600, 2300 RC Leiden, The Netherlands
Department of Molecular Epidemiology, Leiden University Medical Centre, PO Box 9503, 2300 RA Leiden, The Netherlands

Hum Mol Genet. 13, 1755-1762 (2004)

Abstract

Background

Interleukin-10 (IL10) is a cytokine with key regulatory and anti-inflammatory function involved in the pathogenesis of various diseases. Although the large interindividual differences in the production of IL10 have been extensively associated with polymorphisms and haplotypes of the IL10 gene, surprisingly little evidence exists that this variation is actually dictated by IL10 haplotypes.

Methods & Findings

Using the technique of allele-specific transcript quantification, the ratio between two alleles (A and G) of the IL10 gene was characterized in 15 healthy heterozygous individuals. Two groups were identified whereby donors in group 1 exhibited a 1:1 ratio, whereas those in group 2 exhibited a ratio > 1 ($P < 0.0017$). We found that donors heterozygous for haplotype IL10.2 (one of the four ancient IL10 haplotypes) were only prevalent in the group that showed higher allelic expression ratios.

Conclusion

In this study we show that IL10 alleles are indeed differentially transcribed in cells from heterozygous individuals and that IL10 haplotypes dictate production of IL10. These findings show that interindividual differences in IL10 protein levels can be explained at the transcriptional level.

Introduction

Interleukin-10 (IL10) has a key role in regulation of the immune response. Produced by a variety of cell types, including monocytes and B-lymphocytes, it is a potent upregulator of B-cell activation and differentiation (1), but has anti-inflammatory capabilities that can down-regulate TNFa, IL1, IL8 and IFNg (2,3). It has also been shown to be an essential growth factor for Tr1 cells (4).

Studies have shown that striking differences exist between healthy individuals in their ability to produce IL10 both at the constitutive level and following LPS stimulation of wholeblood culture in vitro. The relevance of these interindividual differences in IL10 secretion is supported by several studies. In the case of fertility, a 2-fold more prevalence of an IL10 genotype associated with low production (-2849AA) was observed among 73 women who remained childless as compared with the prevalence among 323 women with normal fecundity (5). In a study of first-degree family members of patients with meningococcal disease, families characterized by high production of IL10 had a 20-fold higher chance of a fatal outcome in comparison with families with low production of IL10 (6). These data along with other studies point to a genetic origin of enhanced IL10 production that is clinically relevant.

Around 50-70% of the interindividual differences in IL10 production can be attributed to genetic factors (6, 7). Polymorphisms in the promoter region of the IL10 locus exhibit functional relevance in a pleiotropy of diseases of infectious, autoimmune or immunosuppressive nature. Eleven promoter single nucleotide polymorphisms (SNPs) and two CA-repeat microsatellites (IL10.G and IL10.R) have been described. This has led to the identification of four common IL10 promoter haplotypes (8). One interesting SNP is -2849A/G (relative to the transcription start site) in the IL10 promoter. Two previous studies from our research showed that carriers of the -2849AA genotype (non-G carriers) produce significantly less IL10 than carriers of the AG and GG genotypes (G carriers) (5,9). The -2849AA genotype together with the IL10.R3 microsatellite tag the IL10.1 haplotype associated with low production of IL10. Interestingly, haplotypes of the IL10 promoter have been implicated in disease outcome as well. In a recent study, Lin et al. demonstrated that an IL10 promoter haplotype is an independent predictor of outcome of graft versus host disease after stem-cell transplantation (10).

The genetic characterization of the IL10 production phenotype was based on results from whole-blood cultures stimulated with LPS. Different amounts of secreted IL10 in response to a constant amount of stimulus showed a high correlation at the mRNA level, while the mRNA half-life between high and low producing donors was the same (8), indicating that secreted protein levels are regulated by differential transcription. However, the intraindividual variation in the amount of IL10 produced by one given individual is about 20% (11), and the associations between IL10 haplotypes and low or high producer phenotype are variable. Although current techniques, using statistic associations and artificial reporter constructs, may be effective in assigning a certain haplotype or a SNP as the marker of low/high IL10 production, we aimed at determining the exact nature of the relationship between IL10 production and cis-acting variations within the IL10 locus.

In general, both copies of human autosomal genes are assumed to be co-dominantly expressed. However, exceptions do arise in the course of normal development or when

promoter utilization or RNA stability is altered (12). Allelic imbalances were previously studied in situations of genomic imprinting when both maternal and paternal alleles are present, but one allele is expressed while the other remains inactive (13). Recent surveys of human and mouse genes suggest that the occurrence of allelic expression differences are more common than previously expected (14,15). Although differential allelic expression may be a potentially important mechanism underlying human phenotypic diversity, no data is currently available showing a direct functional effect of allelic imbalances. Here, we report the first study of genetic variations in the IL10 gene aimed at analysing the intrinsic relationship between genotype and phenotype at the transcriptional level by quantifying the rate of RNA synthesis.

Results

To distinguish between alleles of the IL10 gene, a panel of 58 unrelated, healthy individuals was genotyped for a SNP at nucleotide position +4259 (position relative to the transcription start site) present on exon 5/3'UTR region (Supplementary Material, Fig. S1). Given that this SNP is located within a part of the gene that is transcribed, it is possible to discriminate between allelic transcripts in a +4259 heterozygous individual. Thus, 15 heterozygous unrelated individuals were selected for further analysis.

To validate the allele-specific transcript quantification (12) procedure the intensities of the fragments in DNA mixtures with a known distribution of +4259A and +4259G alleles were measured along with genomic DNA controls. The ratio of the intensity of the bands equals the ratio in the DNA mixtures (Supplementary Material, Fig. S2a) with a high correlation factor (R = 0.999). Genomic DNA yielded an equimolar ratio with a median of 1.00 (25th percentile 0.97-75th percentile 1.035) as expected (Supplementary Material, Fig. S2b). Taken together these results show that any deviations from DNA equimolar ratio can be quantified as each allele acts as an internal control for confounding factors that alter the overall expression of that gene, including differences in tissue preparation, mRNA quality and environmental influences.

Quantification of the respective IL10 +4259A and IL10 +4259G allelic RNA transcripts for 15 individuals revealed that the median ratio of IL10 + 4259 A:G ranged from 0.93 to 1.13 (Fig. 1). In 33% of the individuals (2–6) both IL10 alleles contribute equally to the mRNA pool, whereas 67% showed a preferential expression of allele G or A, respectively (individuals 1 and 7–15). Individual 1 showed preferential expression of allele G significantly different from genomic DNA ($P < 0.0001$). In individuals whose A allele was preferentially expressed (7–15), the ratio of the transcripts as compared with genomic DNA varied from 1.03 to 1.13 ($P < 0.003$). A:G ratios of samples derived from blood collected at a second time-point for individuals 3, 11 and 15 yielded a ratio of 1.01, 1.15 and 1.145, respectively as expected (data not shown), further supporting evidence for allele-specific expression.

We therefore characterized two predominant groups, group 1 which does not show a significant deviation from genomic DNA ratio and group 2 which includes individuals showing preferential expression of allele A by 3–12%. Moreover, upon examination of nuclear families of individuals 13 and 14 (family 1 and family 2, respectively) exhibiting allelic variation in expression (Fig. 2), we observed that higher allele-specific expression in the IL10 gene is a heritable trait.

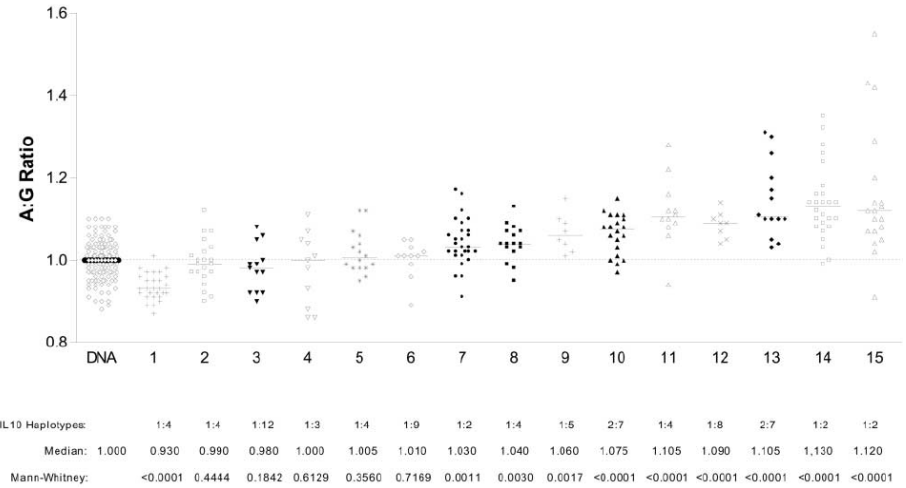


Figure 1. Determination of the relative contribution of the IL10 +4259A and IL10 +4259G alleles in LPS-stimulated whole-blood cultures of 15 healthy individuals. Each dot represents a sample of DNA or cDNA. Data are representative of a minimum of three independent experiments. All individuals (1–15) were subjected to independent whole-blood cultures stimulated with LPS, followed by separate RNA purification and cDNA preparation. P-values refer to significance test for differences (Mann–Whitney U-test) between every individual RNA transcript ratio to genomic DNA ratio (significance $P < 0.05$). Individuals 2–6 yielded a median ratio not significantly different from genomic DNA (Mann–Whitney $P > 0.2$), whereas individuals 1 and 7–15 yielded a ratio of $\neq 1$, significantly different from genomic DNA ($P < 0.0017$). Individually inferred haplotypes for each individual are indicated.

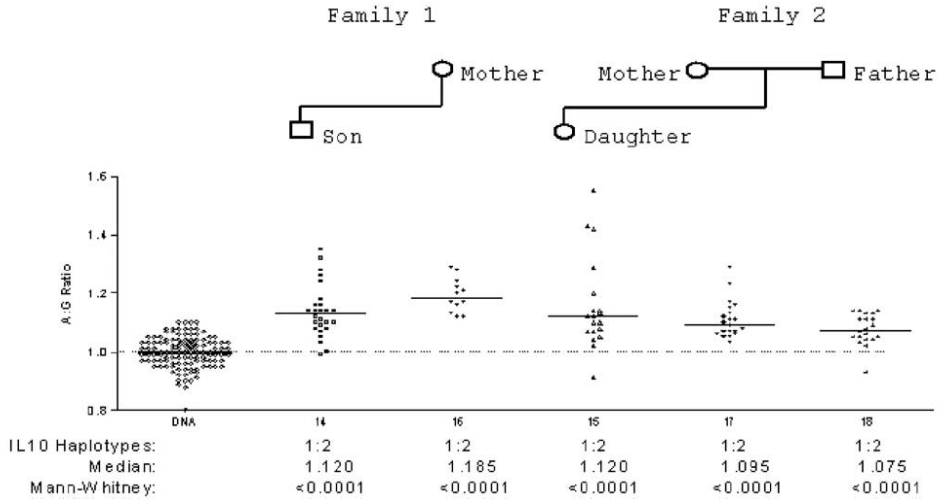


Figure 2. Inheritance pattern of altered IL10 allelic expression. The relative contribution of the IL10 +4259A and IL10 +4259G alleles in whole-blood cultures stimulated with LPS was investigated in two pedigrees (families 1 and 2), as shown, in which the same trait is prevalent implying an underlying inheritance pattern. Each dot represents a sample of DNA or cDNA. Data are representative of a minimum of three independent ASTQ experiments. P-values refer to significance test for differences (Mann–Whitney U-test) between every individual transcription ratio to genomic DNA ratio (significance $P < 0.05$).

Given the extensive associations between IL10 production and IL10 SNP genotypes, we hypothesized that the regulatory determinant could be one of the already characterized IL10 SNPs found in the promoter region (16). Therefore, we analysed if one of the SNPs' genotypes was always homozygous in the 1:1 donors and always heterozygous in the 1: ≠ 1 donors. Table 1 shows that this is not the case. Although none of the SNPs show an individual effect, a combination of alleles may identify which haplotype dictates differential expression. Haplotype frequencies for the population comprising of 58 individuals were subsequently inferred using the estimation maximization (EM) algorithm in SNPHAP (Fig. 3). Two haplotypes were assigned to each heterozygous individual ($P \geq 0.9$ were taken as true haplotypes). The level of linkage disequilibrium (LD) between all SNPs was very high, D' values were close to 1 (data not shown), thus only 13 out of the 29 possible haplotypes were prevalent in the population. IL10.1 to IL10.4 are the four most common haplotypes previously identified (17) and comprise 75% of total chromosomes.

Table 1. IL10 genotypes and haplotypes of 15 unrelated individuals included in the ASTQ analysis and three family members (16–18)

Individual	Sex	Age (y)	RNA A:G (median)	*P*-value	IL10 polymorphisms									IL10 haplotypes	
					A + 4259G	T + 1582C	C − 592A	C − 819T	G − 1082A	A − 1330G	A − 2763C	A − 2849G	A − 3575T		
1	M	45	0.930	<0.0001	AG	CT	CA	CT	AG	AG	AC	AG	TA	4(L)	1(L)
2	M	27	0.990	0.4444	AG	CT	CA	CT	AG	AG	AC	AG	TA	4(L)	1(L)
3	M	28	0.980	0.1842	AG	CC	CC	CC	GG	AA	AA	AA	AA	12(L)	1(L)
4	F	32	1.000	0.6129	AG	CC	CC	CC	GG	AA	AA	AG	AA	3(L)	1(L)
5	F	24	1.005	0.356	AG	CT	CA	CT	AG	AG	AC	AG	TA	4(L)	1(L)
6	M	40	1.010	0.7169	AG	CC	CA	CT	AG	AG	AC	AG	TA	9(L)	1(L)
7	F	26	1.030	0.00011	AG	CC	CC	CC	AG	AG	AC	AG	TA	2(H)	1(L)
8	M	30	1.040	0.003	AG	CT	CA	CT	AG	AG	AC	AG	TA	4(H)	1(L)
9	M	26	1.060	0.0017	AG	CC	CC	CC	GG	AA	AC	AG	TA	5(H)	1(L)
10	F	31	1.075	<0.0001	AG	CC	CC	CC	AG	AG	CC	AG	TA	2(H)	7(L)
11	M	28	1.105	<0.0001	AG	CT	CA	CT	AG	AG	AC	AG	TA	4(H)	1(L)
12	F	35	1.090	<0.0001	AG	CC	CC	CC	GG	AA	AC	AG	AA	8(H)	1(L)
13	F	27	1.105	<0.0001	AG	CC	CC	CC	AG	AG	CC	AG	TA	2(H)	7(L)
14	M	43	1.130	<0.0001	AG	CC	CC	CC	AG	AG	AC	AG	TA	2(H)	1(L)
15	F	32	1.120	<0.0001	AG	CC	CC	CC	AG	AG	AC	AG	TA	2(H)	1(L)
16	F	75	1.185	<0.0001	AG	ND	CC	CC	AG	AG	AC	AG	TA	2(H)	1(L)
17	F	67	1.095	<0.0001	AG	CC	CC	CC	AG	AG	AC	AG	TA	2(H)	1(L)
18	M	71	1.075	<0.0001	AG	CC	CC	CC	AG	AG	AC	AG	TA	2(H)	1(L)

ND, not determined. Each individual's A:G RNA ratio (1–18) is expressed as median. To determine if significant differences exist between RNA ratio and genomic DNA ratio, Mann–Whitney *U*-test was carried out (significance level $P < 0.05$). Nine IL10 SNPs including IL10 +4259 (exon 5/3'UTR), IL10 +1582 (intron 3); proximal promoter SNPs IL10 -592, IL10 -819, IL10 -1082, IL10 -1330 and distal promoter SNPs IL10 -2763, IL10 -2849 and IL10 -3575 were genotyped, and haplotypes for each individual are indicated. L is assigned to the lower IL10 transcribing allele and H to the higher transcribing allele.

Analysis of individual haplotypes in Table 1 shows that the IL10.2 haplotype is only present in the donors that exhibit a >1 ratio ($P < 0.0017$). IL10.2 is therefore proposed to be the 'high' haplotype, responsible for higher IL10 transcription and thus higher IL10 protein levels.

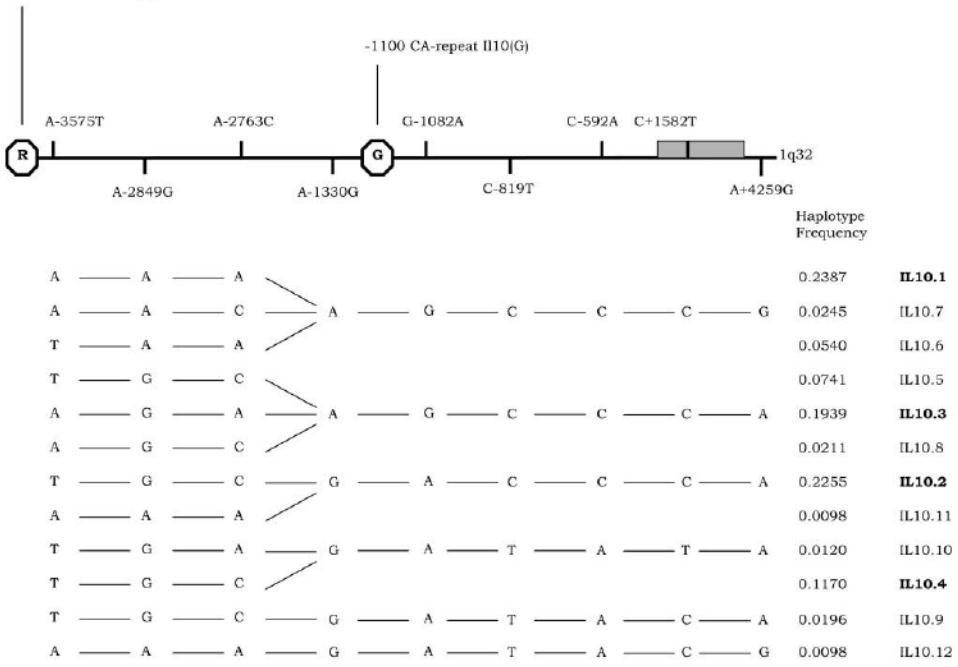


Figure 3. Haplotype frequencies of nine IL10 polymorphisms including seven promoter SNPs in 58 normal Caucasian individuals. SNPs were denoted respective to the transcription start site. In total, 13 haplotypes were inferred using EM algorithm in SNPHAP. Haplotypes IL10.1, 2, 3 and 4 are previously identified ancient haplotypes in the Caucasian population (17) and account for 75% of variation within our population. The proximal region of the IL10 promoter shows less variation than the distal part.

Discussion

IL10 is a crucial player in the immune system. Genetic differences in IL10 production have been associated extensively with the pathology and outcome of diseases. The aim of this study was to determine whether the mechanism driving these associations is differences in allelic expression. We show here for the first time that alleles of the IL10 gene are indeed differentially transcribed in LPS-stimulated whole-blood cultures of heterozygous individuals, a finding which supports the functional existence of transcriptionally distinct IL10 haplotypes in humans and questions the functional relevance of currently known SNPs.

IL10.2 harbours a polymorphism that is associated with increased expression at the mRNA level. This implies that IL10.1 and IL10.7 harbour the SNP associated with lower transcription rate. By extrapolation, high or low phenotype has been assigned to each haplotype and indicated in Table 1. Our data can best be explained if the IL10.4 haplotype consists of two haplotypes, a haplotype associated with high and one associated with low allele-specific IL10 expression.

Our data are in accordance with previously published data on LPS-induced IL10 production. We observed that the -3575A/-2763A haplotype is associated with lower IL10 production (16). In our panel, this combination of -3575A/-2763A haplotype forms part of the larger haplotypes IL10.1, 3 and 12, which are responsible for lower allele-specific IL10 transcription. We can therefore correlate these haplotypes to low IL10 protein levels, providing further evidence that this finding is in line with all data previously obtained. The -3575T/-2849G/-2763C haplotype, which tags IL10.2, IL10.4 and IL10.5, is associated with high IL-10 production, while the -3575A/-2849A/-2763A haplotype, which tags IL10.1 and IL10.12, is associated with low IL-10 production. According to the data in Table 1, both IL10.1 and IL10.12 drive lower transcription of mRNA transcripts, and both IL10.2 and IL10.5 express higher allelic transcriptional levels, further supporting our analysis.

Our finding that the IL10.4 haplotype may either be a low or a highly transcribed haplotype and the occurrence of an individual with ratio <1 suggests the presence of a bidirectional allelic imbalance that could result from allelic heterogeneity of one or more cis-acting regulatory polymorphisms present in the coding, intronic and/or other non-coding regulatory sequences (14,18) or from DNA methylation, histone acetylation and other epigenetic factors (19,20). Alternatively, the marker SNP could be detecting a single causal polymorphism that is actually responsible for the differential expression of IL10 RNA transcripts, but that does not show unidirectional allelic imbalance when it is not in complete LD.

Prime candidates are SNPs that are present in the regulatory region of the IL10 gene. However, our current knowledge of regulatory elements in the human genome is far from comprehensive and, beyond their coding sequences, most genes are not well annotated. As the positions of all regulatory sequences for any given gene are generally unknown, comprehensive direct analyses of such elements are currently impossible. Moreover, the SNPs used in this study span only about 8 kb of the IL10 gene, a region that does not cover the average length of a high LD block in world populations of similar origin (21). It would be interesting to determine the average length of the LD block surrounding the IL10 gene and attempt to identify regulatory sequences within.

In summary, we propose that a common SNP in high but not complete LD with haplotype IL10.2 dictates differences in IL10 expression. This cis-acting variation(s) may alter the functional motifs in the promoter region or work indirectly by modifying the activity of the gene product. However, the magnitude of difference in expression level that is likely to be biologically relevant is unclear. It is plausible that different cell types present in peripheral blood dilute the effect seen, such that investigating specific cell types would yield larger differences in IL10 differential transcription. However, in our system of LPS-induced whole-blood cultures, the main producers of IL10 are monocytes and macrophages, and so far no data exist to support monoallelic expression of monocytes. A more likely alternative, in our view, is that if a particular allele is always transcribing at a faster rate, subtle differences can have an exponential effect resulting in large changes in gene expression patterns. We therefore conclude the mechanism behind higher IL10 gene expression levels is caused by subtle differences in allelic expression regulated by cis-acting polymorphism(s) on distinct IL10 haplotypes.

Materials and Methods

Reagents

All reagents were from Invitrogen, Life Technologies, Gaithersburg, MD, USA unless otherwise specified.

Subjects/DNA isolation

Blood samples were collected from 58 unrelated healthy Caucasian volunteers. DNA was isolated using a SOC-lysis and proteinase-K treatment of peripheral blood cells, followed by a phenol–chloroform extraction. From every heterozygous IL10 + 4259A/G (rs3024498) individual, two additional heparin-blood tubes were used to perform LPS-whole-blood stimulations, as described below.

Genotyping of IL10-polymorphisms (position relative to transcription start site)

The IL10 -3575A/T (rs1800890), -2849A/G, -2763A/C, -1082A/G (rs1800896) and -819A/C (rs1800871) polymorphisms were typed as previously described (22). An amount of 25 ng genomic DNA was used per polymerase chain reaction (PCR) for all SNPs. -592A/C (rs1800872) and -1330A/G (rs1800893) are in complete LD (R^2 = 1) with -819A/C and -1082A/G, respectively. All SNPs were annotated relative to the transcription start site according to the public reference sequence U16720.

The PCR for -592A/C genotyping was carried out using 0.6 U Taq polymerase with the buffer supplied by the manufacturer (Applied Biosystems), 2.5 mM MgCl2, 0.25 mM of each dNTP, 100 µg/ml bovine serum albumin and 0.25 µM of each primer in a 30 µl reaction. Sense primers were 5'-CTC AGTTAGCACTGGTGTAC-3' and antisense 5'-TGTTCCTA GGTCACAGTGAC-3'. DNA was denatured at 94°C for 5 min and cycling conditions were set at 94°C for 30 s, 60°C for 30 s and 72°C for 30 s for 35 cycles and a final cycle of extension at 72°C for 7 min in a thermocycler (Applied Biosystems 9700). An aliquot of 10 µl of the PCR product was digested for 2 h with 1.5 U of Rsa I (MBI Fermentas) in a total volume of 25 µl at 37°C, yielding 480 bp for the C allele and 240 bp for the A allele.

The PCR conditions for the +4259A/G SNP were the same as for the -592A/C and were run for 40 cycles. Primers were sense 5'-ACTGAGCTTCTCTGTGAACG-3' and antisense 5'-AATAACAAGCTGGCCACAGC-3'. 12.5 µl PCR product was digested with 2.5 U per tube Fnu4HI (New England Biolabs) in a total volume of 50 µl yielding 416 bp for the A allele and 284 + 132 bp for the G allele.

The conditions for the +1582T/C (rs1554286) SNP PCR were set at 1 min 94°C, 1 min 60°C and 1 min 72°C for 40 cycles with sense primer 5'-CTACGGCGCTGTGTAAGTAGCAGATCAGA*T-3' (modified nucleotides are marked) and antisense primer 5'-AACCCACAAATGACTCACAAAT-3'. An aliquot of 10 µl PCR product was digested with 2.5 U of Bgl II in a total volume of 25 µl yielding 157 + 105 bp for the T allele and 130 + 105 + 27 bp for the C allele.

Quality control

Data for each SNP were reviewed independently to verify their quality. We considered each SNP to be validated and correctly genotyped for all individuals only if the genotype data met a series of strict criteria, including minimum signal intensity specifications and unambiguous genotype reading ability. No errors were detected in 10% of the samples that were randomly genotyped again.

LD mapping and haplotyping across the IL10 gene

GOLD (23) was used to determine pairwise LD between all SNPs. Haplotype frequencies were inferred using SNPHAP (http://www-gene.cimr.cam.ac.uk/clayton/software/). Crosschecking of the inferred haplotypes using available family data (data not shown) revealed that all haplotypes were accurate in areas of high LD (D' close to 1).

LPS-whole-blood cultures/RNA isolation

In total, 15 heterozygous unrelated individuals were selected for ASTQ on the basis of their heterozygosity with respect to the IL10 +4259 SNP. Approximately 20 ml of blood was collected at one time-point from each of 18 heterozygous IL10 + 4259A/G donors (15 unrelated and three family members, Table 1). Blood was collected from three individuals at a second time-point and an independent set of experiments was carried out.

For each donor, five wells containing 3 ml of blood was immediately diluted once in 3 ml of an aliquot consisting of endotoxin-free RPMI 1640 medium (Gibco) with 10% fetal bovine serum (Bodinco BV), 5% pen-strep (Bio-Whittaker) and 30 mg/ml LPS from Salmonella typhosa (Sigma). As a control, one additional well was subjected to the same conditions in the absence of LPS. These aliquots were incubated for 17 h at 37°C with 5% CO2.

RNA isolation

The samples were harvested, red blood cells lysed, and centrifuged at 600g. Each cell pellet was washed once with 25 ml PBS, taken up in 1 ml RNA-Bee (Campro Scientific) and stored at -20°C for at least 18 h. These samples were thawed on ice and 100 µl chloroform was added. After the samples were vigorously shaken for 20 s, and incubated on ice for 60 min, they were centrifuged (13 200 rpm, 4°C) for 15 min. The upper aqueous phase was pipetted out into new tubes. An equal amount of isopropanol was added to each tube followed by incubation on ice for 15 mins. RNA was precipitated by centrifugation (13 200 rpm, 4°C) for 15 min. The supernatant was discarded and the pellets were washed with 500 µl 70% ethanol. Samples were vortexed and centrifuged (13 200 rpm, 4°C) for 5 min. Supernatant was discarded and pellets were dried in air for 5–10 min. RNA was solubilised in 40–50 µl DEPC-treated water by incubation in a 55°C water bath for 30 min.

DNAse-treatment/RT-PCR

RNA samples were treated with DNAse I, Amp. grade. Each DNAse-treated RNA sample was split into four equal volumes, each containing 2 µg RNA. Two fractions were reverse-transcribed into cDNA in a 40 µl volume at 37°C for 1 h using 400 U of M-MLV, 5x first-strand buffer, 0.1 M dithiotreitol, rRNAsin (Promega) 100 mM dNTP mix and either 1 µg of oligo $dT_{12\text{-}18}$, or 1 µg of random primers. The remaining two fractions were used as negative controls. The reaction was inactivated by incubation at 95°C for 10 min. In total, each blood collection from each individual yielded six RNA samples which were reverse transcribed into 12 cDNAs. All cDNA samples were tested several times in independent ASTQ experiments as outlined below. Samples that were not stimulated with LPS were excluded from the analysis as the intensities obtained were too low to be quantified.

Allele-specific transcript quantification

This method is essentially the same as previously described (12) with the exception of the use of a marker (IL10 +4259) which is present in exon 5/3'UTR region of the IL10 gene, thus allowing the post-transcriptional analysis of the expressed alleles.

cDNA samples were amplified by PCR with the IL10 +4259A/G PCR. An aliquot of 20 µl amplicon was mixed with 4 µCi [α-^{32}P]-dCTP (Amersham Biosciences) and 0.5 U Taq polymerase, followed by one PCR cycle with conditions set at 94°C for 5 min, 94°C for 1 min, 60°C 1 min, 72°C for 1 min and 72°C for 7 min. An aliquot of 15 µl labelled amplicon was digested with 2.5 U Fnu4HI in a total volume of 50 µl and incubated at 37°C for 2 h. All samples were ethanol-salt precipitated, dissolved and separated on a 5% polyacrylamide sequence gel. After electrophoresis, the gel was dried for 3 h in a preheated vacuum gel-dryer and incubated for 7 days in a Phosphor imager (Molecular Dynamics) cassette at room temperature. The exposed screen was scanned and computer analysed (ImageQuant).

Ratios of +4259A/+4259G transcripts were obtained by calculating the pixel values of digestion fragments of allele A (416 bp) versus allele G (284 bp). As the respective products are not of the same size, a correction factor was applied based on the respective GC content of each restriction product: GC content allele A = 168 bp; GC content allele G = 108 bp; Correction factor = 168/108 = 1.56 was applied to each cDNA ratio. As the values for these two bands reflect the synthetic products from a RT-PCR amplification of the same RNA preparation, controlling for variations in reaction conditions was unnecessary.

Genomic DNA controls

Each gel consisted of cDNA samples along with genomic DNA controls from six to eight independent individuals. As allele A and allele G are codominantly expressed at the genomic DNA level, A:G DNA ratio from heterozygous IL10 + 4259 individuals should be 1:1. The reliability of the cDNA data was determined by the average of the DNA ratios obtained. In each gel, RNA ratios were subjected to a correction factor equivalent to genomic DNA ratio to circumvent overestimation of differences. Gels with a genomic DNA ratio <0.95 and >1.05 were excluded from further analysis.

Statistics

Data were analysed using Graphpad prism and did not follow a normal distribution. The median value for RNA A:G ratio was calculated for each individual (Table 1) and a Mann–Whitney U-test was performed to detect significant differences between RNA ratios and genomic DNA ratio ($P < 0.05$).

References

(1) Rousset, F., Garcia, E., Defrance, T., Peronne, C., Vezzio, N., Hsu, D.H., Kastelein, R., Moore, K.W. and Banchereau, J. (1992) Interleukin 10 is a potent growth and differentiation factor for activated human B lymphocytes. Proc. Natl Acad. Sci. USA, 89, 1890–1893.

(2) de Waal, M.R., Abrams, J., Bennett, B., Figdor, C.G. and de Vries, J.E. (1991) Interleukin 10 (IL-10) inhibits cytokine synthesis by human monocytes: an autoregulatory role of IL-10 produced by monocytes.J. Exp. Med., 174, 1209–1220.

(3) Katsikis, P.D., Chu, C.Q., Brennan, F.M., Maini, R.N. and Feldmann, M. (1994) Immunoregulatory role of interleukin 10 in rheumatoid arthritis. J. Exp. Med., 179, 1517–1527.

(4) Kemper, C., Chan, A.C., Green, J.M., Brett, K.A., Murphy, K.M. and Atkinson, J.P. (2003) Activation of human CD4+ cells with CD3 and CD46 induces a T-regulatory cell 1 phenotype. Nature, 421, 388–392.

(5) Westendorp, R.G., van Dunne, F.M., Kirkwood, T.B., Helmerhorst, F.M. and Huizinga, T.W. (2001) Optimizing human fertility and survival. Nat. Med., 7, 873.

(6) Westendorp, R.G., Langermans, J.A., Huizinga, T.W., Elouali, A.H., Verweij, C.L., Boomsma, D.I., Vandenbroucke, J.P. and Vandenbrouke, J.P. (1997) Genetic influence on cytokine production and fatal meningococcal disease. Lancet, 349, 170–173.

(7) Reuss, E., Fimmers, R., Kruger, A., Becker, C., Rittner, C. and Hohler, T. (2002) Differential regulation of interleukin-10 production by genetic and environmental factors—a twin study. Genes Immun., 3, 407–413.

(8) Eskdale, J., Gallagher, G., Verweij, C.L., Keijsers, V., Westendorp, R.G. and Huizinga, T.W. (1998) Interleukin 10 secretion in relation to human IL-10 locus haplotypes. Proc. Natl Acad. Sci. USA, 95, 9465–9470.

(9) de Jong, B.A., Westendorp, R.G., Eskdale, J., Uitdehaag, B.M. and Huizinga, T.W. (2002) Frequency of functional interleukin-10 promoter polymorphism is different between relapse-onset and primary progressive multiple sclerosis. Hum. Immunol., 63, 281–285.

(10) Lin, M.T., Storer, B., Martin, P.J., Tseng, L.H., Gooley, T., Chen, P.J. and Hansen, J.A. (2003) Relation of an interleukin-10 promoter polymorphism to graft-versus-host disease and survival after hematopoietic-cell transplantation. N. Engl. J. Med., 349, 2201–2210.

(11) van der Linden, M.W., Huizinga, T.W., Stoeken, D.J., Sturk, A. and Westendorp, R.G. (1998) Determination of tumour necrosis factor-alpha and interleukin-10 production in a whole blood stimulation system: assessment of laboratory error and individual variation. J. Immunol. Methods, 218, 63–71.

(12) Kaijzel, E.L., Bayley, J.P., van Krugten, M.V., Smith, L., van de, L.P., Bakker, A.M., Breedveld, F.C., Huizinga, T.W. and Verweij, C.L. (2001) Allele-specific quantification of tumor necrosis factor alpha (TNF) transcription and the role of promoter polymorphisms in rheumatoid arthritis patients and healthy individuals. Genes Immun., 2, 135–144.

(13) Reik, W. and Walter, J. (2001) Genomic imprinting: parental influence on the genome. Nat. Rev. Genet., 2, 21–32.

(14) Cowles, C.R., Hirschhorn, J.N., Altshuler, D. and Lander, E.S. (2002) Detection of regulatory variation in mouse genes. Nat. Genet., 32, 432–437.

(15) Bray, N.J., Buckland, P.R., Owen, M.J. and O'Donovan, M.C. (2003) Cis-acting variation in the expression of a high proportion of genes in human brain. Hum. Genet., 113, 149–153.

(16) Gibson, A.W., Edberg, J.C., Wu, J., Westendorp, R.G., Huizinga, T.W. and Kimberly, R.P. (2001) Novel single nucleotide polymorphisms in the distal IL-10 promoter affect IL-10 production and enhance the risk of systemic lupus erythematosus. J. Immunol., 166, 3915–3922.

(17) Eskdale, J., Keijsers, V., Huizinga, T. and Gallagher, G. (1999) Microsatellite alleles and single nucleotide polymorphisms (SNP) combine to form four major haplotype families at the human interleukin-10 (IL-10) locus. Genes Immun., 1, 151–155.

(18) Martone, R., Euskirchen, G., Bertone, P., Hartman, S., Royce, T.E., Luscombe, N.M., Rinn, J.L., Nelson, F.K., Miller, P., Gerstein, M. et al. (2003) Distribution of NF-kappaB-binding sites across human chromosome 22. Proc. Natl Acad. Sci. USA, 100, 12247–12252.

(19) Dillon, N. and Festenstein, R. (2002) Unravelling heterochromatin: competition between positive and negative factors regulates accessibility. Trends Genet., 18, 252–258.

(20) Jaenisch, R. and Bird, A. (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat. Genet., 33 (suppl.), 245–254.

(21) Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M. et al. (2002) The structure of haplotype blocks in the human genome. Science, 296, 2225–2229.

(22) Moraes, M.O., Santos, A.R., Schonkeren, J.J., Vanderborght, P.R., Ottenhoff, T.H., Moraes, M.E., Moraes, J.R., Sampaio, E.P., Sarno, E.N. and Huizinga, T.W. (2003) Interleukin-10 promoter haplotypes are differently distributed in the Brazilian versus the Dutch population. Immunogenetics, 54, 896–899.

(23) Abecasis, G.R. and Cookson, W.O. (2000) GOLD—graphical overview of linkage disequilibrium. Bioinformatics., 16, 182–183. 1762 Human Molecular Genetics, 2004, Vol. 13, No. 16

Chapter 8

Interleukin 10: a new risk marker for the development of restenosis after percutaneous coronary intervention

P.S. Monraats
F.A.S. Kurreeman
D. Pons
V.D.K.D. Sewgobind
F.R. de Vries
A.H. Zwinderman
M.P.M. de Maat
P.A. Doevendans
R.J. de Winter
R.A. Tio
J. Waltenberger
T.W.J. Huizinga
D. Eefting
P.H.A. Quax
R.R. Frants
A. van der Laarse
E.E. van der Wall
J.W. Jukema

Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands
Interuniversity Cardiology Institute of The Netherlands (ICIN), Utrecht, The Netherlands
Department of Rheumatology, Leiden University Medical Center, Leiden, TheNetherlands
Department of Medical Statistics, Academic Medical Center, Amsterdam, The Netherlands
Department of Hematology, Erasmus University Medical Center, Rotterdam, The Netherlands
Department of Cardiology, University Medical Center, Utrecht,The Netherlands
Department of Cardiology, Academic Medical Center, Amsterdam, The Netherlands
Department of Cardiology, University Medical Center Groningen, Groningen, The Netherlands
Department of Cardiology, Academic Hospital Maastricht, Maastricht, The Netherlands
Gaubius Laboratory, TNO Quality of Life, Leiden, The Netherlands
Department of Vascular Surgery,Leiden University Medical Center, Leiden, The Netherlands
Department of Human Genetics, Center for Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands

Genes and Immunity (2007) 8, 44–50

Abstract

Genetic factors appear to be important in the process of restenosis after percutaneous coronary intervention (PCI), as well as in inflammation, a pivotal factor in restenosis. An important mediator in the inflammatory response is interleukin (IL)-10. Our aim was to study whether genetic variants in IL-10 predispose to the risk of restenosis. The GENetic DEterminants of Restenosis (GENDER) study included 3104 patients treated with successful PCI. Target vessel revascularization (TVR) was chosen as primary end point.

Genotyping of the -2849G/A, -1082G/A, -592C/A and +4259A/G polymorphisms of the IL-10 gene was performed by MassArray platform. After adjusting for clinical variables, three polymorphisms significantly increased the risk of restenosis (-2849AA: relative risk (RR), 1.7, 95% confidence interval (CI), 1.2–2.5; -1082AA: RR, 1.4, 95% CI, 1.1–1.8 and +4259GG: RR, 2.0, 95% CI, 1.4–2.8). To further exclude possible involvement of neighboring genes due to LD in the IL-10 locus, additional polymorphisms were genotyped. The results reveal that association of the IL-10 gene with restenosis is independent of flanking genes.

Our findings demonstrate that IL-10 is associated with restenosis and therefore support the hypothesis that anti-inflammatory genes also may be involved in developing restenosis. Furthermore, they may provide a new targeting gene for drug-eluting stents.

Introduction

An important limitation of the treatment of atherosclerotic lesions by percutaneous coronary intervention (PCI) is the occurrence of restenosis(1,2). Inflammatory responsiveness, resulting in neointima formation, plays a pivotal role in the development of restenosis (3–5). Several inflammatory genes have already been reported to be associated with the development of restenosis (6,7). However, little is known about the involvement of anti-inflammatory cytokines, although they seem logic candidate genes in the process of restenosis. Interleukin 10 (IL-10) is one of these anti-inflammatory genes. It is an important immunosuppressor cytokine, involved in the regulation of many aspects of immune responses. Its effects are directed mainly against functions of mononuclear cells, T lymphocytes and polymorphonuclear leukocytes. Furthermore, IL-10 plays a role in inhibition of cell adhesion molecules, monocyte chemotactic protein-1, tissue factor, fibrinogen, matrix metalloproteinase-9, T-lymphocyte granulocyte–macrophage colony-stimulation factor, inducible nitric oxide synthase and smooth muscle cell proliferation (8–10). Several of these factors have been demonstrated to be involved in the restenotic process (5,11). The interindividual difference among individuals in their ability to produce IL-10 appears to have a genetic origin (12–14). The heritability of the endotoxininduced IL-10 production has been estimated to be 74% in studies on monozygotic or dizygotic twins and nonrelated individuals (15). The gene encoding IL-10 contains variable sites (polymorphisms and microsatellite markers) that have previously been associated with the level of IL-10 produced, indicating that they may be associated with different responsiveness to regulatory signals.

The aim of this study was to assess whether four different functional polymorphisms (three in the promoter, one in the 3'UTR) in the IL-10 gene are related to the risk of developing restenosis after PCI.

Results

A total of 3146 patients had a complete follow-up (99.3%) with a median duration of 9.6 months (interquartile range 3.9 months). Out of 3146 patients, 42 had an event in the first 30 days. These patients were excluded from further analysis, according to the protocol. Baseline characteristics of the population are shown in Table 1. Genotyping was successful in 2874 patients for the -2849G/A polymorphism, in 2740 patients for the -1082G/A polymorphism, in 2873 patients for the -592C/A polymorphism and in 2865 patients for the +4259A/G polymorphism. Allele frequencies were 0.72/0.28, 0.51/0.49, 0.76/0.24 and 0.73/0.27, respectively. The results of the remaining patients are missing owing to the lack of DNA or inconclusive genotyping. Patients who could not be genotyped did not differ in any characteristic from those who could be genotyped. The distributions of the genotypes are shown in Table 2. All polymorphisms showed no significant deviation from Hardy–Weinberg equilibrium ($P>0.05$), except for -1082G/A ($P=0.01$).

Table 1 Demographic, clinical and lesion characteristics of 3104 patients with and without TVR

	Patients with TVR (n = 304)	*Patients without TVR (n = 2800)*	*Total (n = 3104)*
Age (years)	61.7 ± 10.1	62.2 ± 10.8	62.1 ± 10.7
BMI (kg/m²)	26.9 ± 3.7	27.0 ± 3.9	27.0 ± 3.9
Male sex	220 (72.4%)	1996 (71.3%)	2216 (71.4%)
Diabetes	63 (20.7%)	390 (13.9%)	453 (14.6%)
Hypercholesterolemia	188 (61.8%)	1702 (60.8%)	1890 (60.9%)
Hypertension	138 (45.4%)	1121 (40.0%)	1259 (40.6%)
Current smoker	62 (20.4%)	700 (25.0%)	762 (24.5%)
Family history of MI	121 (39.8%)	977 (34.9%)	1098 (35.4%)
Previous MI	109 (35.9%)	1130 (40.4%)	1239 (39.9%)
Previous PCI	64 (21.1%)	493 (17.6%)	557 (17.9%)
Previous CABG	36 (11.8%)	340 (12.1%)	376 (12.1%)
Stable angina	198 (65.1%)	1881 (67.2%)	2079 (67.0%)
Multivessel disease	148 (48.7%)	1284 (45.9%)	1432 (46.1%)
Peripheral vessel disease	12 (3.9%)	92 (3.3%)	104 (3.4%)
Lipid-lowering medication	171 (56.3%)	1516 (54.1%)	1687 (54.3%)
Restenotic lesions	27 (8.9%)	181 (6.5%)	208 (6.7%)
Total occlusions	56 (18.4%)	372 (13.3%)	428 (13.8%)
Type C lesion	94 (30.9%)	708 (25.3%)	802 (25.8%)
Proximal LAD	70 (23.0%)	619 (22.1%)	689 (22.2%)
RCX	75 (24.7%)	764 (27.3%)	839 (27.0%)
Residual stenosis >20%	51 (16.8%)	299 (10.7%)	350 (11.3%)
Stent length (mm)	10.3 (0–82)	13.0 (0–93)	15 (0–146)

Abbreviations: BMI, body mass index; CABG, coronary artery bypass grafting; LAD, left anterior descending branch of the left coronary artery; MI, myocardial infarction; PCI, percutaneous coronary intervention; RCX, circumflex branch of the left coronary artery; VR, target vessel revascularization. Age is mean ±s.d.; other variables are percentage of patients.

Of the 3104 patients, 304 (9.8%) patients underwent target vessel revascularization (TVR) during follow-up. Fifty-one (1.6%) patients died and 22 (0.7%) suffered from myocardial infarction (MI). After univariate analysis, -2849AA, -1082AA and +4259GG genotypes of the IL-10 gene increased the risk for TVR significantly (P=0.005, P=0.03 and P=0.001,respectively) (Table 2).

The effect of each polymorphism was adjusted for patient and intervention-related characteristics that were previously found to be related to TVR risk including diabetes, hypertension, stenting, residual stenosis >20%, current smoking and total occlusion, as well as age and gender. This analysis showed a significant association for the same three polymorphisms that were significantly associated with TVR in the univariate analysis. Furthermore, diabetes, stenting and total occlusion were significantly associated with TVR (Table 3a). Multivariable Cox regression analysis was performed, in which all four IL-10 polymorphisms were included. Subsequently, we performed multivariable Cox regression analysis in which we included all four IL-10 polymorphisms, and adjusted for the same clinical risk factors and selected in a backward stepwise manner the polymorphisms that were independently associated with TVR risk. Polymorphisms were removed from the model when

their P-value was 40.10. A significant association was found for the IL-10 +4259GG genotype (P=0.001), implying that this polymorphism is an independent risk factor for TVR. Exclusion of the -1082G/A polymorphism, which was not in complete HW equilibrium, did not alter the outcome.

As we found a strong correlation between the IL-10 polymorphisms and stenting, we stratified patients to a stented and a non-stented population. The use of intracoronary stents was carried out at the discretion of the operator. The stented population consisted of 2309 patients, of who 203 (8.8%) patients had to undergo a TVR. Both the stented population and the non-stented population, consisting of 795 patients, demonstrated a significant association for the same three polymorphisms as described earlier when adjusted for clinical variables. Multivariable analysis including all polymorphisms demonstrated the +4259GG genotype of the IL-10 gene to be associated with TVR. Results of this analysis are shown in Table 3b and 3c. Furthermore, they show that the effect of the +4259GG genotype of the IL-10 gene was more pronounced in the non-stented population.

Table 2 Distributions of the polymorphisms, including the univariate analysis of investigated polymorphisms in association with TVR

Polymorphisms	*Number of cases and controls genotyped* (N%)	*Minor allele frequency*	*Best-fitting genetic model*	*TVR rate for the different genotypes* (%)	*P-value**
−2849G/A		0.28	Recessive		0.005
GG/GA	2659 (92.5)			9.1	
AA	215 (7.5)			14.9	
−1082G/A		0.49	Recessive		0.03
GG/GA	2040 (74.5)			8.7	
AA	700 (25.5)			11.4	
−592C/A		0.24	Additive		0.42
CC	1693 (58.9)			10.0	
CA	1008 (35.1)			9.0	
AA	172 (16.0)			9.3	
+4259A/G		0.27	Recessive		0.001
AA/AG	2643 (92.3)			9.1	
GG	222 (7.7)			16.7	

Abbreviation: TVR, target vessel revascularization.
*P-value determined by the Cox proportional regression model.

Table 3 Relative risks (RR) of the univariate and multivariable analysis of the IL-10 polymorphisms in association with TVR for the
(a) total population (N = 3104); (b) stented population (N = 2309) and (c) non-stented population (N = 795)

	Raw RR (95% CI)	*Adjusted for clinical variables RR (95% CI)*	*Multivariable analysis, including clinical variables and all four polymorphisms*
(a)			
Diabetes	—	1.5 (1.1–2.1)	1.5 (1.1–2.1)
Total occlusion	—	1.4 (1.0–1.9)	1.4 (1.01–1.9)
Hypertension	—	NS	NS
Stenting	—	0.7 (0.5–0.9)	0.7 (0.5–0.9)
Restenosis > 20%	—	NS	NS
IL-10 –592C/A	0.9 (0.8–1.1)	NS	NS
IL-10 –2849G/A	1.7 (1.2–2.4)	1.7 (1.2–2.5)	NS
IL-10 –1082G/A	1.4 (1.04–1.8)	1.4 (1.1–1.8)	NS
IL-10 +4259A/G	1.9 (1.4–2.7)	2.0 (1.4–2.8)	2.0 (1.4–2.9)
(b)			
Diabetes	—	1.5 (1.1–2.1)	1.6 (1.1–2.4)
Total occlusion	—	1.4 (1.01–1.9)	NS
Hypertension	—	NS	NS
Restenosis > 20%	—	NS	NS
IL-10 –592C/A	1.0 (0.8–1.3)	0.9 (0.7–1.1)	NS
IL-10 –2849G/A	1.2 (0.7–2.0)	1.7 (1.2–2.5)	NS
IL-10 –1082G/A	1.2 (0.9–1.7)	1.4 (1.1–1.8)	NS
IL-10 +4259A/G	1.5 (0.9–2.4)	2.0 (1.4–2.8)	1.6 (1.0–2.6)
(c)			
Diabetes	—	1.5 (1.1–2.1)	NS
Total occlusion	—	1.4 (1.01–1.9)	NS
Hypertension	—	NS	NS
Restenosis > 20%	—	NS	NS
IL-10 –592C/A	0.8 (0.5–1.1)	0.9 (0.7–1.1)	NS
IL-10 –2849G/A	2.7 (1.6–4.7)	1.7 (1.2–2.5)	NS
IL-10 –1082G/A	1.6 (1.02–2.5)	1.4 (1.1–1.8)	NS
IL-10 +4259A/G	2.9 (1.7–5.0)	2.0 (1.4–2.8)	3.1 (1.8–5.4)

Abbreviations: 95% CI = 95% confidence interval; IL, interleukin; NS, not significant; RR = relative risk; TVR, target vessel revascularization.

A high level of linkage disequilibrium exists between the polymorphisms in the IL-10 gene. In order to determine whether the three significant associated polymorphisms with TVR are acting synergistically to confer risk to TVR, we performed two haplotype analyses. We did not include the -592C/A polymorphism in the haplotype analysis, as it did not show a significant association with TVR. Furthermore, the R^2 value for this polymorphism is much lower (pairwise linkage disequilibrium (LD) between -592 and +4259 has R^2=0.12).

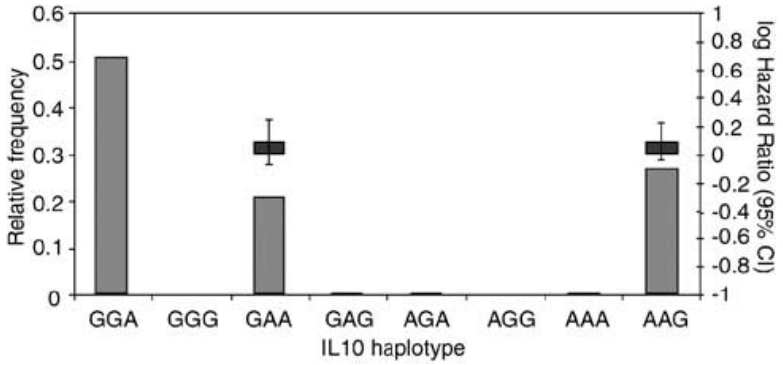


Figure 1 IL-10 and haplotypes. Haplotypes are presented in the following order of polymorphisms: -2849G/A, -1082G/A and +4259A/G. Of the eight possible haplotypes, only three haplotypes had relative frequency >1%, presented by the red bars. The blue bars present log hazard ratios. Differences were not statistically significant (P>0.05).

We first performed a combined analysis with -2849G/A, -1082G/A and +4259A/G. Of the eight possible haplotypes, only seven were observed with relative frequency >0, and only three haplotypes had relative frequency >1%, namely 'GGA', 'GAA' and 'AAG' (Figure 1, red bars). The frequency of the GGA haplotype in patients with TVR was 0.47 compared to 0.51 for patients without TVR. For the GAA haplotype, the haplotype frequency was 0.22 for patients with TVR and 0.21 for patients without TVR. Furthermore, patients with TVR had a frequency of the AAG haplotype of 0.30 compared to 0.27 for patient without TVR. Compared to the wild-type haplotype ('GGA'), the log hazard ratios of the other haplotypes were 0.094 and 0.098, respectively (Figure 1, blue bars). Differences were not statistically significant (P>0.05). Furthermore, as the -1082G/A polymorphism was not in HW equilibrium and the-2849G/A and the +4259A/G are in strong linkage disequilibrium (D' = 0.96, R^2=0.92 pairwise LD as calculated by haploview), we performed separate haplotype analysis for the -2849G/A and the +4259A/G polymorphisms. The combination of these two polymorphisms did not give an additive effect (data not shown).

IL-10 is located in a cluster of IL-10 family genes on chromosome 1q32. As the extent of LD surrounding the IL-10 gene has not as yet been definitively characterized, we investigated whether the effect seen from the examined IL-10 polymorphisms could in fact be the result of LD with a neighboring gene. To this end, we genotyped polymorphisms in mitogen-activated protein kinase-activated protein kinase 2 (MAPKAPK2), IL-10, IL-19, IL-20 and IL-24 in a healthy population of unrelated Caucasian individuals (*N*=60). Haploview was used to assess

pairwise LD between polymorphisms of genes. Figure 2 reveals that the region spanning 217 kb is broken down into four blocks with a solid spine of LD. MAPKAPK2, IL-19, IL-20 and IL-24 seem to be independent of the IL-10 gene itself. IL-19, IL-20 and IL-24 form combined blocks, indicating shared inheritance patterns of IL-20 with both IL-19 and IL-24. These data are in line with the haplotype blocks as defined by HAPMAP for a Caucasian population of European descent. In our study, this finding excludes the presence of *cis* effects of the IL-10 homologs or MAPKAPK2 on the IL-10 gene owing to ancestral recombination and reveals that the association of the IL-10 gene with restenosis is independent of flanking genes.

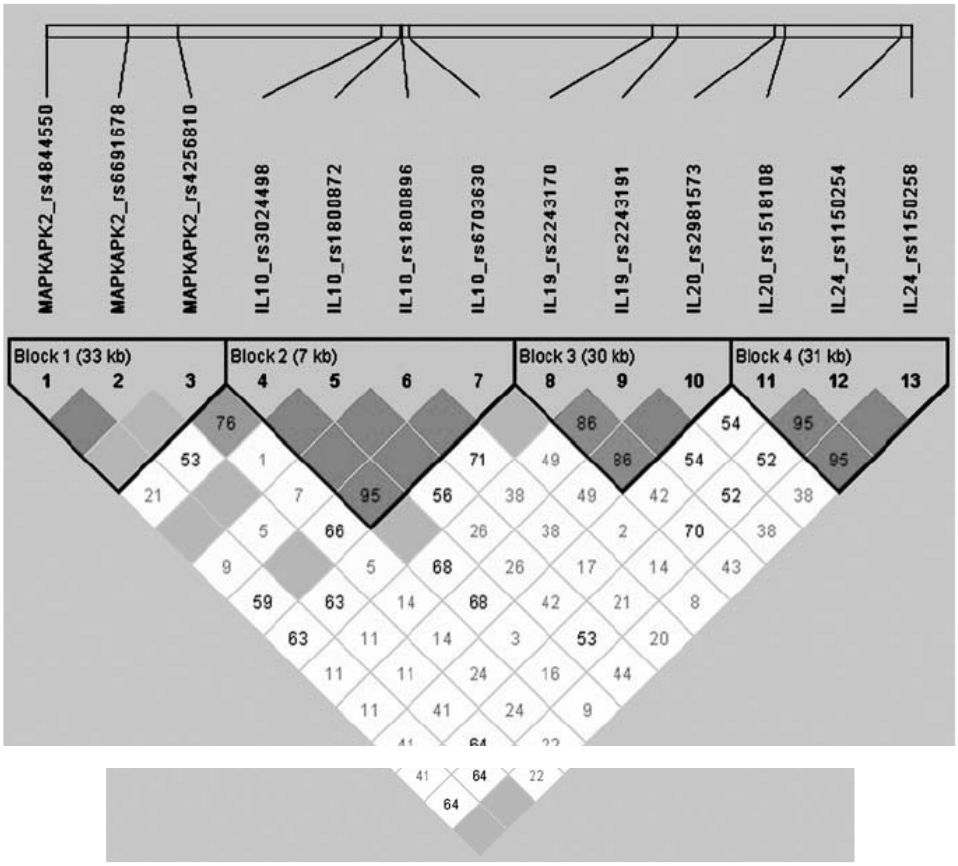


Figure 2 Haplotype map of the IL-10 locus. The haplotype map is made by Haploview version 3.2 software. The basis of this graphical representation is pairwise linkage disequilibrium (LD) between polymorphisms. D' values (%) are shown in the boxes. In the event of D' = 100%, the boxes are empty. The region consists of four recombination blocks that encompass a 217 kb region between MAPKAPK2 and IL-24. Block 1 comprises the gene MAPKAPK2, block 2 comprises the best-known promoter polymorphisms of IL-10 and spans the whole IL-10 gene (-592C/A

Discussion

Inflammation is known to play a pivotal role in the development of restenosis after PCI (4,5). Many inflammatory genes have already been investigated in relation to restenosis. However, the role of anti-inflammatory genes in restenosis is not fully understood. Therefore, we investigated in our large prospective GENDER study the effect of four different polymorphisms of the IL-10 gene in relation to restenosis, defined by TVR in our study. Different cell types, including human monocytes and T cells, produce IL-10. IL-10 inhibits the production of pro-inflammatory cytokines, including tumor necrosis factor alpha (TNFα) (16). Furthermore, it interferes with the production of various chemokines (17). Treatment of IL-1- activated human endothelial cells with IL-10 results in lower surface densities of intercellular adhesion molecule-1 and vascular-cell adhesion molecule-1 and reduced leukocyte adhesivity (18). In addition, IL-10 enhances the production of an IL-1 receptor antagonist that has anti-inflammatory activity directed against the effects of IL-1 (19).

After adjustment for clinical variables, we found the -2849AA, -1082AA and +4259GG genotypes of the IL-10 gene to be highly associated with restenosis after PCI. After inclusion of all four polymorphisms and clinical factors previously associated with TVR, only the +4259GG genotype showed a significant association. Possibly, the +4259A/G polymorphism has an effect on restenosis by influencing the mRNA stability.

Haplotype analysis did not provide any additional information with regards to risk of TVR. As the -1082G/A polymorphism was not in complete HW equilibrium, and as we cannot be certain that this is due to selection of the population, or to genotype errors, we also performed multivariable analysis excluding this polymorphism. Exclusion of this polymorphism did not significantly alter the results. With the current chosen polymorphisms, we are able to characterize about 70% of the four common haplotypes in the IL-10 gene. However, we cannot exclude the possibility that less common haplotypes, characterized by further tagging polymorphisms (as provided by the HAPMAP project, www.hapmap.org), could provide more information with regards to the risk of restenosis.

The results of our study reveal that association of the IL-10 gene with restenosis is independent of flanking genes, as *cis*-acting variations in neighboring genes are unlikely to play a role given that there are clear recombination points around the IL-10 gene, breaking the LD with neighboring genes. These findings are in line with the results of the HAPMAP project in a Caucasian population of European descent.

Some studies found a relation between TNFα and IL-10 plasma levels (20). As we previously investigated the role of several polymorphisms in the TNFα gene (11) we analyzed whether those polymorphisms had an effect on our results. However, we did not find a significant association between the -238G/A and the -1031T/C polymorphisms of the TNFα gene and the four IL-10 polymorphisms we examined in this study.

The functional effect of the polymorphisms we examined has been described previously. Koss *et al* demonstrated that the A allele in the IL-10 promoter region at position -1082 was associated with decreased IL-10 production as measured by enzyme-linked immunosorbent assay (ELISA) in lipopolysaccharide (LPS)-stimulated whole blood in Crohn's disease patients and healthy controls (P=0.005, P=0.015, respectively) (13). Furthermore, several studies have demonstrated that carriers of the -2849AA genotype have significantly lower IL-10 responsiveness upon stimulation with endotoxin (12,21). In some individuals, allele G of the +4259 polymorphism produces less IL-10 transcripts as compared to the A allele (14), which points towards an allele-specific genetic regulation of protein levels of IL- 10. As we showed an increase in the risk of developing restenosis for the -1082AA, -2849AA and +4259GG genotypes, these data corroborate our hypothesis that lower levels of IL-10 may increase the risk of developing restenosis. However, the relevant stimulus and therefore the relevant transcription factor activating the IL-10 gene in the case of restenosis is largely unknown. Different transcription factors are known to either positively or negatively regulate the transcription of the IL-10 gene (22, 23). The current functional data available on the regulation of the IL-10 gene is thus a mere indication of plausible scenarios that may or may not provide explanations in the case of restenosis.

Eefting *et al*. studied the involvement of IL-10 in neointima formation in a hypercholesterolemic mouse model of cuff-induced stenosis of the femoral artery by IL-10 knocking-out or overexpression procedures. Knocking out IL-10, in hypercholesterolemic ApoE*3-Leiden mice, resulted in a significant 1.9-fold increase of neointima surface as compared to ApoE*3- Leiden IL-$10^{+/+}$ littermates (P=0.02). Conversely, a marked 45% inhibition on cuff-induced neointima formation was obtained after IL-10 overexpression (P=0.02) (24).

Another study has examined the effect of recombinant human IL-10 (rhuIL-10) on intimal growth, after angioplasty or stent implantation, in hypercholesterolemic rabbits (25). The main findings of their study were that systemic administration of the anti-inflammatory cytokine rhuIL-10 successfully inhibits intimal hyperplasia after balloon injury or stent implantation in hypercholesterolemic rabbits. This protective effect is associated with a major inhibition of IL-1b release by circulating leukocytes and reduced infiltration of the arterial wall by activated macrophages. RhuIL-10 has no apparent effect on lipid metabolism and no systemic toxicity in their animal model (25). It is therefore possible that an IL-10-releasing stent may contribute to lowering the development of restenosis.

Study limitations

Circulating protein levels were not assessed in the present study. However, we believe that basal (pre-PCI) lasma levels of the gene product will not reflect the genetically determined differences in IL-10 increase after a trauma such as PCI. Moreover, local differences in response (in the vessel wall at the place of PCI) may not be reflected systemically. In humans, it is impossible to measure gene products locally in the acute phase of treatment or the following days, and several months later the causal trigger has probably already disappeared.

Furthermore, the lack of data on the effect of the presently used drug-eluting stents on TVR and its relationship with the IL-10 polymorphisms is a limitation of our study. Another potential limitation is that we examined TVR as our primary end point instead of angiographic outcomes, such as late loss. However, in clinical practice, clinical restenosis is an end point much more valuable than angiographic restenosis. Finally, as our study was conducted in a sample of Caucasian patients, extrapolation of the data to other ethnic groups should be done with great caution.

Conclusions

The present study shows that the -2849AA, -1082AA and +4259GG genotypes of the gene coding for the anti-inflammatory cytokine IL-10 are a risk marker for the development of restenosis. Further investigation in other populations as well as the fine mapping of the IL-10 gene will provide further insight into the precise role of IL-10 in restenosis. Based on our findings, screening patients for this genotype can lead to a better stratification of patients at increased risk for restenosis and thereby provide indications for improving individual tailor-made treatment, as it may be a new target point for drug-eluting stents. The results of this study lend support to the broader hypothesis that genetic programming of the inflammatory response plays a significant role in the development of restenosis. Given the explorative nature of this analysis, our results need to be reproduced in other studies.

Materials & Methods

GENDER project

Study design

The present study population has been described previously (26). In brief, the GENDER project was designed to study the association between various gene polymorphisms and clinical restenosis. Patients eligible for inclusion were treated successfully for stable angina, non-ST-elevation acute coronary syndromes or silent ischemia by PCI in four out of 13 referral centers for interventional cardiology in the Netherlands. Patients treated for acute ST elevation MI were excluded. Also excluded from analysis were patients suffering from events occurring within 1 month after PCI, as these events were attributable predominantly to sub-acute stent thrombosis or occluding dissections, rather than to restenosis.

PCI procedure

Experienced operators, using a radial or femoral approach, performed standard angioplasty and stent placement. Before the procedure, patients received aspirin 300mg and heparin 7500 IU. The use of intracoronary stents and additional medication, such as glycoprotein IIb/IIIa inhibitors, were carried out at discretion of the operator. In case a stent was implanted, patients received either ticlopidin or clopidogrel for at least 1 month following the procedure depending on local practice. During the study, no drug-eluting stents were used.

Follow-up and study end points

Follow-up lasted for at least 9 months, except when a coronary event occurred. Patients were either seen in outpatient clinics or contacted by telephone. TVR, either by PCI or coronary artery bypass grafting (CABG), was designated the primary end point, as it is considered most relevant by regulatory agencies. An independent clinical events committee evaluated the clinical events.

The study protocol meets the criteria of the Declaration of Helsinki and was approved by the Medical Ethics Committees of each participating institution. Written informed consent was obtained from all participating patients before the PCI procedure.

Genetic methodology

Blood was collected in ethylene diaminetetraacetate (EDTA) tubes at baseline and genomic DNA was extracted following standard procedures. In this population, we determined genotypes of the following polymorphisms in the IL-10 gene: -2849G/A (rs6703630), -1082G/A (rs1800896), -592C/A (rs1800872) and +4259A/G (rs3024498). These polymorphisms were selected from literature, and databases on the web, criteria used were frequency of the rare allele and a possible functional effect (8,10,13,14). To assay these polymorphisms, we used a MassArray platform according to manufacturer's protocols. Two multiplex assays were designed using Assay designer software (Sequenom, Hamburg, Germany). After polymerase chain reaction (PCR), a primer extension reaction was performed to introduce mass-differences

between alleles and, after removing salts by adding a resin, ~15 nl of the product was spotted onto a target chip with 384 patches containing matrix. Mass differences were detected using an Autoflex (Bruker, Wormer, Netherlands) matrix-assisted laser desorption\ionization time-of-flight mass spectrometry (MALDI-TOF) and genotypes were assigned real-time using Typer 3.1 software (Sequenom). As quality controls, 5–10% of the samples were genotyped in duplicate. No inconsistencies were observed. Positive and negative controls uniquely distributed in each 384 wells plate were also consistent. Cluster plots were made of the signals from the low and the high mass allele. Two independent researchers carried out scoring. Disagreements or vaguely positioned dots produced by Genotyper 3.0 (Sequenom) as well as all wells that had 50% or more failed SNPs were excluded from analysis.

Statistical analysis

Deviations of the genotype distribution from that expected for a population in HW equilibrium was tested using the χ^2 test with one degree of freedom. Allele frequencies were determined by gene counting, the 95% CIs of the allele frequencies were calculated from sample allele frequencies, based on the approximation of the binominal and normal distributions in large sample sizes.

Continuousvariables are expressed as mean± standard deviation and were compared by means of the unpaired, two-sided *t*-test. Discrete variables are expressed as counts or percentages and were compared with the χ^2. In the first stage, the association between the IL-10 polymorphisms and TVR was assessed using the Cox proportional regression model under a co-dominant genetic model. No adjustments for covariates were performed at this stage, so that we could assess their possible involvement in the causal pathway.

All polymorphisms were also assessed using dominant and recessive models, and the model with the lowest Akaike information criteria was used in multivariable regression analysis (27). Multivariable regression analysis of the TVR risk was performed with all IL-10 polymorphisms, using a stepwise backward selection algorithm. In the final step, clinical variables associated with TVR, also including age and gender, were entered into the regression model. The IL-10 polymorphisms were combined into haplotypes and the effect of haplotypes on restenosis risk was estimated according to the methods developed by Tanck *et al* (28). Evaluation of the neighboring genes of the IL-10 locus was performed by genotyping polymorphisms in the MAPKAPK2, IL-19, IL-20 and IL-24 genes in a panel of healthy individuals (*N=60*).

Haploview software (http://www.broad.mit.edu/mpg/haploview) was used to perform LD calculations. Haplotype blocks were assigned under the algorithm of solid spine of LD as provided by the software. Statistical analysis was carried out using SPSS 12.0 for Windows (SPSS Inc., Chicago, IL, USA). A P-value <0.05 was considered statistically significant.

References

(1) Mehilli J, Kastrati A, Bollwein H, Dibra A, Schuhlen H, Dirschinger J et al. Gender and restenosis after coronary artery stenting. Eur Heart J 2003; 24: 1523–1530.

(2) Schomig A, Kastrati A, Elezi S, Schuhlen H, Dirschinger J, Dannegger F et al. Bimodal distribution of angiographic measures of restenosis six months after coronary stent placement. Circulation 1997; 96: 3880–3887.

(3) Sharif F, Daly K, Crowley J, O'Brien T. Current status of catheter- and stent-based gene therapy. Cardiovasc Res 2004; 64: 208–216.

(4) Agema WR, Jukema JW, Pimstone SN, Kastelein JJ. Genetic aspects of restenosis after percutaneous coronary interventions: towards more tailored therapy. Eur Heart J 2001; 22: 2058–2074.

(5) Monraats PS, Pires NM, Agema WR, Zwinderman AH, Schepers A, De Maat MP et al. Genetic inflammatory factors predict restenosis after percutaneous coronary interventions. Circulation 2005; 112: 2417–2425.

(6) De Maat MP, Jukema JW, Ye S, Zwinderman AH, Moghaddam PH, Beekman M et al. Effect of the stromelysin-1 promoter on efficacy of pravastatin in coronary atherosclerosis and restenosis. Am J Cardiol 1999; 83: 852–856.

(7) Kastrati A, Koch W, Berger PB, Mehilli J, Stephenson K, Neumann FJ et al. Protective role against restenosis from an interleukin-1 receptor antagonist gene polymorphism in patients treated with coronary stenting. J Am Coll Cardiol 2000; 36: 2168–2173.

(8) Lowe PR, Galley HF, Abdel-Fattah A, Webster NR. Influence of interleukin-10 polymorphisms on interleukin-10 expression and survival in critically ill patients. Crit Care Med 2003; 31: 34–38.

(9) Mazighi M, Pelle A, Gonzalez W, Mtairag EM, Philippe M, Henin D et al. IL-10 inhibits vascular smooth muscle cell activation in vitro and in vivo. Am J Physiol Heart Circ Physiol 2004; 287: H866–H871.

(10) Suarez A, Castro P, Alonso R, Mozo L, Gutierrez C. Interindividual variations in constitutive interleukin-10 messenger RNA and protein levels and their association with genetic polymorphisms. Transplantation 2003; 75: 711–717.

(11) Monraats PS, Pires NM, Schepers A, Agema WR, Boesten LS, de Vries MR et al. Tumor necrosis factor-alpha plays an important role in restenosis development. FASEB J 2005; 19: 1998–2004.

(12) de Jong BA, Westendorp RG, Eskdale J, Uitdehaag BM, Huizinga TW. Frequency of functional interleukin-10 promoter polymorphism is different between relapse-onset and primary progressive multiple sclerosis. Hum Immunol 2002; 63: 281–285.

(13) Koss K, Satsangi J, Fanning GC, Welsh KI, Jewell DP. Cytokine (TNF alpha, LT alpha and IL-10) polymorphisms in inflammatory bowel diseases and normal controls: differential effects on production and allele frequencies. Genes Immun 2000; 1:185–190.

(14) Kurreeman FA, Schonkeren JJ, Heijmans BT, Toes RE, Huizinga TW. Transcription of the IL10 gene reveals allelespecific regulation at the mRNA level. Hum Mol Genet 2004; 13: 1755–1762.

(15) Westendorp RG, Langermans JA, Huizinga TW, Verweij CL, Sturk A. Genetic influence on cytokine production in meningococcal disease. Lancet 1997; 349: 1912–1913. 16 Cassatella MA, Meda L, Bonora S, Ceska M, Constantin G. Interleukin 10 (IL-10) inhibits the release of proinflammatory cytokines from human polymorphonuclear leukocytes. Evidence for an autocrine role of tumor necrosis factor and IL-1 beta in mediating the production of IL-8 triggered by lipopolysaccharide. J Exp Med 1993; 178: 2207–2211.

(16) Kasama T, Strieter RM, Lukacs NW, Burdick MD, Kunkel SL. Regulation of neutrophil-derived chemokine expression by IL-10. J Immunol 1994; 152: 3559–3569.

(17) Krakauer T. IL-10 inhibits the adhesion of leukocytic cells to IL-1-activated human endothelial cells. Immunol Lett 1995; 45: 61–65.

(18) Cassatella MA, Meda L, Gasperini S, Calzetti F, Bonora S. Interleukin 10 (IL-10) upregulates IL-1 receptor antagonist production from lipopolysaccharide-stimulated human polymorphonuclear leukocytes by delaying mRNA degradation. J Exp Med 1994; 179: 1695–1699.

(19) Booy R, Nadel S, Hibberd M, Levin M, Newport MJ. Genetic influence on cytokine production in meningococcal disease. Lancet 1997; 349: 1176.

(20) Westendorp RG, van Dunne FM, Kirkwood TB, Helmerhorst FM, Huizinga TW. Optimizing human fertility and survival. Nat Med 2001; 7: 873.

(21) Liu YW, Chen CC, Tseng HP, Chang WC. Lipopolysaccharideinduced transcriptional activation of interleukin-10 is mediated by MAPK- and NF-kappaB-induced CCAAT/enhancer-binding protein delta in mouse macrophages. Cell Signal 2006; 18: 1492–1500.

(22) Riemann M, Endres R, Liptay S, Pfeffer K, Schmid RM. The IkappaB protein Bcl-3 negatively regulates transcription of the IL-10 gene in macrophages. J Immunol 2005; 175: 3560–3568.

(23) Eefting D, Schepers A, de Vries MR, Pires NMM, Grimbergen JM, Lagerweij T et al. The effect of interleukin-10 knock-out and overexpression on neointima formation in hypercholesterolemic ApoE3Leiden mice. Manuscript in progress.

(24) Feldman LJ, Aguirre L, Ziol M, Bridou JP, Nevo N, Michel JB et al. Interleukin-10 inhibits intimal hyperplasia after angioplasty or stent implantation in hypercholesterolemic rabbits. Circulation 2000; 101: 908–916.

(25) Agema WRP, Monraats PS, Zwinderman AH, de Winter RJ, Tio RA, Doevendans PAFM et al. Current PTCA practice and clinical outcomes in The Netherlands: the real world in the pre-drug-eluting stent era. Eur Heart J 2004; 25: 1163–1170.

(26) Li W, Nyholt DR. Marker selection by Akaike information criterion and Bayesian information criterion. Genet Epidemiol 2001; 21 (Suppl 1): S272–S277.

(27) Tanck MW, Klerkx AH, Jukema JW, De Knijff P, Kastelein JJ, Zwinderman AH. Estimation of multilocus haplotype effects using weighted penalised log-likelihood: analysis of five sequence variations at the cholesteryl ester transfer protein gene locus. Ann Hum Genet 2003; 67: 175–184.

Chapter 9

Genetic Variants at the Interleukin 10 locus do not associate with Rheumatoid Arthritis Disease Outcomes

F.A.S. Kurreeman
D. van der Woude
M. Thabet
K. Verpoort
J. Houwing-Duistermaat
A.H.M. van der Helm-van Mil
R.E.M. Toes
T.W.J. Huizinga

Department of Rheumatology, Leiden University Medical Center, Leiden, Netherlands
Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, The Netherlands

In preparation

Abstract

Objective

Interleukin 10 has been previously identified as a genetic risk factor involved in the severity of joint destruction in Rheumatoid Arthritis (RA). However many studies provide conflicting results when investigating RA phenotypes such as radiographic damage or clinical remission. The aim of this study was to determine whether tagging SNPs in the *IL10* locus are associated with RA disease outcomes in a well-defined clinical cohort.

Methods

RA patients enrolled in the Leiden Early Arthritis Clinic were genotyped for six tagging SNPs in the *IL10* gene (n=594). Yearly radiographs were scored using the Sharp-van der Heijde method. Sustained DMARD-free remission (defined as no swollen joints for at least 1 year after discontinuing all DMARDs) was assessed in all patients. The extent of joint destruction between genotype as well as haplotype groups was compared using Kruskal-Wallis tests at each time point for a follow-up period of 5 years after inclusion. The association between genotype, haplotype groups and remission was investigated using Cox regression analysis.

Results

Patients did not show significant differences in the rate of joint damage and remission with respect to the genotypes and haplotypes they harbour for any of the six tag SNPs. No significant differences were also observed when stratifying for the presence or absence of anti-citrullinated protein antibodies (ACPA).

Conclusion

IL10 polymorphisms most likely do not predispose to a higher rate of joint damage or clinical drug free remission in RA patients irrespective of their ACPA status. These data suggest that *IL10* polymorphisms do not play a major role in RA disease outcomes.

Introduction

Rheumatoid arthritis (RA) is one of the most common autoimmune diseases, affecting ~1% of the population worldwide. Environmental as well as genetic factors are thought to play an important role in both the onset and the progression of the disease (2-5). Recent years have seen large progress in the identification of genetic risk factors involved in RA. To date, fifteen loci identified by genome-wide as well as candidate gene approaches have been implicated in RA. These loci include the HLA region with specific alleles conferring the largest effect size, PTPN22 now widely replicated as an autoimmune locus and many other loci with modest effect sizes(6;7).

Such progress is however as yet unmet in dissecting genetic loci involved in mediating either clinical remission or inversely joint destruction in these patients. Genetic markers associated with these disease-classifying phenotypes are clearly of interest not only to generate more insight into the disease process but possibly also to serve as surrogate predictors of disease course. The sole genetic locus consistently correlated to a more severe radiographic damage is the HLA-shared epitope alleles in the presence of autoantibodies (8). While other loci have been reported, unequivocal replication in large well characterised datasets remains to be obtained.

The *interleukin-10* (*IL10*) locus on chromosome 1q32 has been previously suggested to be associated with RA as well as with a higher rate of joint damage in RA patients(9-12). Since 50-60% of the variation in the level of this protein is likely to be attributed to genetic risk factors, *IL10* promoter polymorphisms have been the prime candidates investigated(13-15). These polymorphisms as well as *IL10* haplotypes have been widely correlated to IL10 mRNA and protein levels(16;17). Very recently, a novel inducible transcript with an extended 5'UTR has been described following lipopolysaccharide (LPS) stimulation of mononuclear cells(18). However extensive analysis of all available genetic variation has so far not been performed in most studies.

Given the important role of this cytokine in mediating immunosuppressive as well as immunostimulatory effects, it remains an important candidate to examine in RA(19). However, the well-powered genome-wide scan performed by the Wellcome Trust Case-Control Consortium consisting of 2000 patients and 3000 healthy individuals has not reported the association of this region with RA. We observed only 54% coverage of this region using SNPs in a 200kb region around *IL10* using a pairwise R^2 ≥0.8 and a coverage of 70% using a pairwise R^2 ≥0.5)(20). More recently a meta-analysis of all published genome-wide association studies also did not identify the *IL10* region as a susceptibility locus for RA(7). These data suggest that the *IL10* region is unlikely to play a major role in the predisposition to RA as such. However, it remains likely that genetic variants at this locus play a role in the various phenotypes of RA characterized by either joint damage or remission. The identification of genetic markers correlating with disease phenotypes remains a crucial question that may not only result in a better understanding of the disease process but also in improving treatment strategies for patients with a worse disease prognosis.

In this study, we have aimed to investigate the correlation of *IL10* polymorphisms to the rate of joint damage and remission in RA patients by performing a comprehensive effort to capture most of the informative markers in the *IL10* region known to date. Combined with the availability of a large number of RA patients and longstanding follow-up, we have determined the association of these SNPs with the progression of radiological joint destruction and remission in a prospective cohort of Early Arthritis patients (EAC).

Patients & Methods

594 RA patients from the population-based arthritis inception cohort (EAC)(21) were included in this study. All RA patients used fulfilled the American College of Rheumatology 1987 revised classification criteria for RA and were selected from a large prospective Early Arthritis Clinic (EAC) cohort that was started in 1993 at the Department of Rheumatology of the Leiden University Medical Center(22). Patients were referred to the EAC by general practitioners in the western part of The Netherlands when arthritis was suspected. The Leiden EAC is the only referral center in an area of ~400,000 inhabitants. Inclusion took place when arthritis was confirmed at physical examination and the symptom duration was less than 2 years. Written informed consent was obtained from all participants. The study and its protocols were approved by the appropriate local institutional review boards. All patients included in the study were of Caucasian origin based on self-reported ethnicity. Five hundred ninety four RA patients who were consecutively included and had DNA available were genotyped and the patient characteristics are summarized in Table 1. The number of patients with available radiographs varied per time point.

At inclusion, blood samples were taken from every patient for routine diagnostic laboratory screening and stored to determine ACPA at a later time point. ACPA were measured by enzyme-linked immunosorbent assay (ELISA) (Immunoscan RA Mark 2; Euro-Diagnostica, Arnhem, The Netherlands). The cutoff level for positivity was set at 25 arbitrary units, according to the manufacturer's instructions. Follow-up visits were performed on a yearly basis and included radiographs of hands and feet. Since the start of the EAC the treatment strategies have changed and four different strategies were applied depending on the inclusion period. Patients included between 1993 and 1995 were treated initially with analgesics and subsequently with chloroquine or salazopyrine if they had persistent active disease (delayed treatment)(23). From 1996 to 1998 RA patients were promptly treated with either chloroquine or salazopyrine (early treatment) (22;23). From 1998 to 2002 patients were promptly treated with either salazopyrine or methotrexate (early treatment) and patients included in 2002 or later were promptly treated with either salazopyrine or methotrexate combined with treatment adjustments based on the disease activity (early and disease activity based treatment).

Genotyping of IL10 polymorphisms and Haplotype block determination

43 SNPs spanning the region around the *IL10* gene encompassing 226kb were genotyped in 57 unrelated healthy Caucasian individuals. Out of the 43 SNPs genotyped, 2 SNPs were not polymorphic (rs2945417 and rs17015763) and one SNP had a frequency ≤1% (rs4845140). Additional details are further described in Table S1. Haploview was used to determine the graphical representation of LD(24). 6 Tagging SNPs were selected and further genotyped in DNA from 594 RA patients. These 6 tagging SNPs captured 86% of all the alleles within the 17kb block with an R^2 threshold of 0.8. Rs3024505 was not tagged using either pairwise LD measures or 2 to 3 marker haplotypes.

Standard quality control for genotyping was carried out. 10% of all genotypes were repeated and <1% errors detected. The success rate greatly exceeded 95% and varied per SNP. More specifically the success rate was 96% for rs6667202, 98% for rs6676671, 100% for rs6703630, 98% for rs6693899, 99.7% for rs1800896 and 99% for rs1800871. All SNP genotypes were in Hardy-Weinberg equilibrium.

Haplotypes per individual were inferred using plink(25). The most likely haplotypes were included in the analysis. The number of individuals with haplotype posterior probability <80% is 26 indicating that uncertainty is minimal for the larger subset of RA patients.

Radiographs

Radiographs of hands and feet were taken on consecutive years starting at baseline and were scored according to the Sharp-van der Heijde method(26). Compared to another frequently used scoring method, the Sharp-van der Heijde method is the most sensitive on the individual patient level in early RA(27). To encompass a reliable sample size during follow-up, radiographic follow-up data were restricted to a maximum of 5 years. The number of available radiographs varied per time-point and was 509 at baseline which declined to 440 after 1 year of follow-up, 401, 340, 283, and 260 after 2 to 5 years of follow-up respectively. Due to the study design (an inception cohort) not all patients achieved a similar duration of follow-up. The cut-off after 5 years of follow-up was applied according to the criteria that at least 10 patients should be available per genotype group. All radiographs were scored by one experienced scorer who was blinded with respect to the patient's autoantibody status, treatment, clinical outcome and genotyping results. Scoring was performed with known time order, which is more sensitive to change, compared to scoring with unknown time sequence(28). From the total number of scored radiographs, 499 radiographs were rescored by the same reader, consisting of 149 baseline radiographs and 350 radiographs during follow-up belonging to 60 randomly selected RA-patients. Reliability of radiograph scoring was calculated and intraclass-observer correlation coefficients (ICC) were 0.91 for all scored radiographs, 0.84 for baseline radiographs and 0.97 for the radiographic progression rate (van der Linden *et al*, Manuscript in press, A&R).

Remission

Sustained DMARD-free remission was defined according to the following three criteria: 1) no current use of DMARDs, 2) no swollen joints and 3) classification as DMARD-free remission by the patient's rheumatologist (Van der Woude *et al*, Manuscript in press at A&R). Corticosteroids were considered to be equivalent to DMARDs for the present study, while NSAIDs did not qualify as DMARDs. Patients had to fulfill all three criteria in order to be diagnosed with remission. To ensure that remission was not temporary, but rather sustained and long-lasting, the absence of swollen joints had to have been observed by a rheumatologist for at least one year after discontinuation of when DMARD-therapy. Patients with remission were discharged from the outpatient clinic if the absence of joint swelling had been observed for at least one year after the discontinuation of DMARDs. Most patients who achieved remission were followed-up longer than the minimum requirement of one year; the median time of observation after discontinuation of DMARDs in the absence of swollen joints was 2.5 years. Patients who had a recurrence of their arthritis after discharge, could easily return to the Leiden University Medical Center, the only referral center for Rheumatology in a health care region of approximately 400.000 inhabitants. The frequency of relapse was recorded and patients with relapse (n=6) were included in the non-remission group. In contrast to previous analyses in which patients who presented to the EAC after 2003 were excluded, the current analysis contained all RA patients for whom *IL10*-genotyping information was available (n=594). This resulted in a slightly lower prevalence of remission (13.2%) than described previously (15%) (Van der Woude et al, manuscript in press at A&R).

Statistical Analysis

The association between genotypes and radiographic scoring data was analyzed using SPSS version 16.0 (SPSS Inc., Chicago, IL). As radiographic data were not normally distributed, the raw data on the Sharp-van der Heijde scores are presented using medians. The Kruskal-Wallis nonparametric test was used to compare median scores across genotype groups. Since six SNPs were evaluated a Bonferroni correction for multiple testing was applied and the p-value for significance was set at $p<0.004$.

Summary statistics were generated to investigate the prevalence of the different *IL10* genotypes and haplotypes in remission and non-remission patients. To avoid skewing of the results due to the difference in follow-up time, the present analysis used data from the first 10 years of follow-up for all patients. To take into account the difference in follow-up times among patients, analyses were performed by Cox regression analysis, after verification that the proportional hazards assumption was satisfied. In Cox regression models the dependent variable is the "time-to-event", which consisted of the time to remission for the remission patients, and the time to last follow-up (with a maximum of 10 years) for the non-remission patients. The time of remission was defined as the date at which DMARDs were discontinued due to remission. The analysis was also performed with a later date defined as the time of remission (date described above plus one year), which led to similar results. This indicates that the results were stable regardless of the exact date used to define remission.

Results

Determination of LD in the IL10 locus

In order to determine the extent of LD in the *IL10* locus, 40 SNPs in the *IL10* region spanning 226kb were genotyped in 57 healthy unrelated Dutch Caucasian individuals. Under the algorithm of Gabriel *et al*, one haplotype block was inferred (Figure 1)(1). This block is restricted to 17 kb encompassing the coding region of the *IL10* gene (5kb) as well as 5'and 3' untranslated regions. Since we are primarily interested in identifying cis-variants predisposing to increased joint damage in RA patients, this block represented the most interesting region to look into. We used Tagger as implemented in Haploview to determine the tagging SNPs that provided the most information with regards to the haplotype variation(29). 6 tagging SNPs were chosen on the basis of pairwise comparisons with $R^2 \leq 0.8$ and a multi-marker LOD threshold of 3. These SNPs captured 93% of all the alleles in the 17kb block. This block was comparable to the haplotype block in HapMap CEU (Figure S1).

Figure 1. 17 kb LD Block around IL10 according to Gabriel et al(1). Paiwise linkage disequilibrium (LD) values (R^2) are indicated. Strength of LD is positively correlated to color intensity.

Association of tagging SNPs and haplotypes with radiological joint damage in RA

The study cohort consisted of 594 patients, whose characteristics at baseline are detailed in Table 1. None of the genotype distributions deviated from Hardy Weinberg equilibrium. Because Sharp-van der Heijde scores did not fit a normal distribution, we compared the median scores across the three genotypes per time point for all six SNPs using the Kruskall-Wallis non-parametric test.

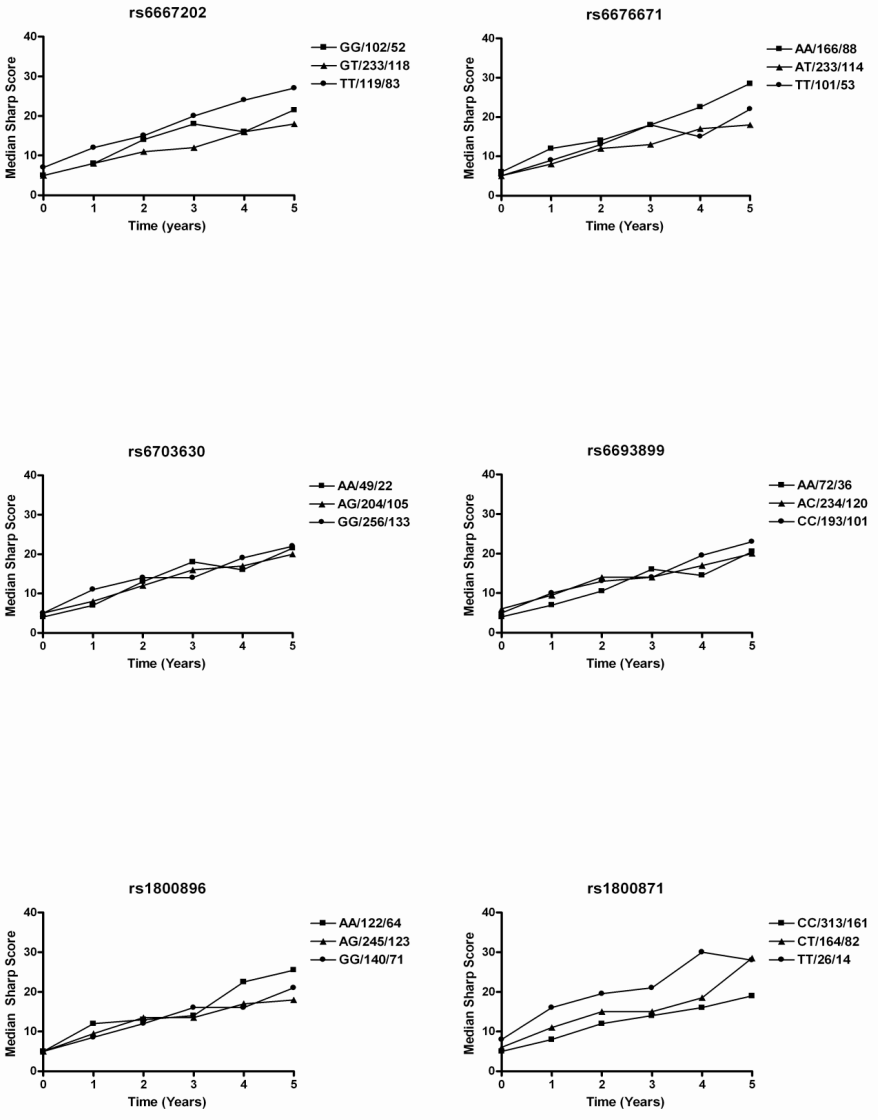
Table 1. Patient characteristics from the Leiden Early arthritis cohort.

Patient Characterisitcs	N=594
Age at inclusion, mean years (± SD)	56.6 (±15.5)
Female N (%)	401 (67.7)
ACPA positive	266 (56.1)
IgM-RF positive	341 (58.3)

Age and gender data were available for all RA patients. IgM-RF (rheumatoid factor) data was available from 585 RA patients. ACPA (anti-citrullinated protein antibodies) data was available from 474 RA patients

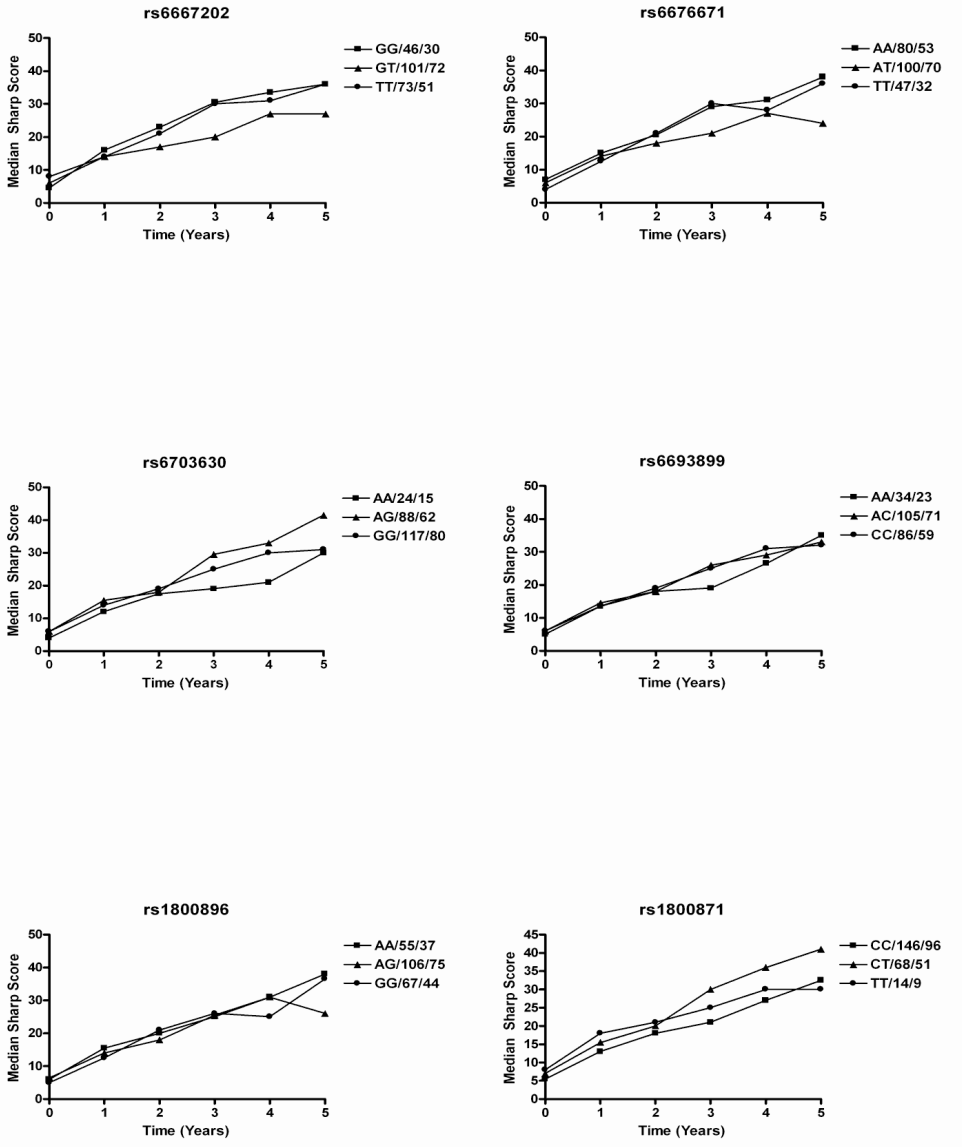
None of the *IL10* tagging SNPs showed a correlation with joint damage in time (Figure 2). Since more severe radiographic damage correlates with the presence of rheumatoid factor and anti-citrullinated protein antibodies (ACPA), we have stratified our data according to the presence or absence of ACPA. No correlation was observed between genotype and joint damage in either stratum (Figure 3 and 4). While the data indicate that patients without ACPA and with rs1800871 TT genotype may have a higher rate of joint destruction, the medians are based solely on 7 individuals at baseline and 4 individuals after 5 years of follow-up.

Figure 2. Median Sharp-van der Heijde Scores for six IL10 tagging SNPs per genotype in all RA patients



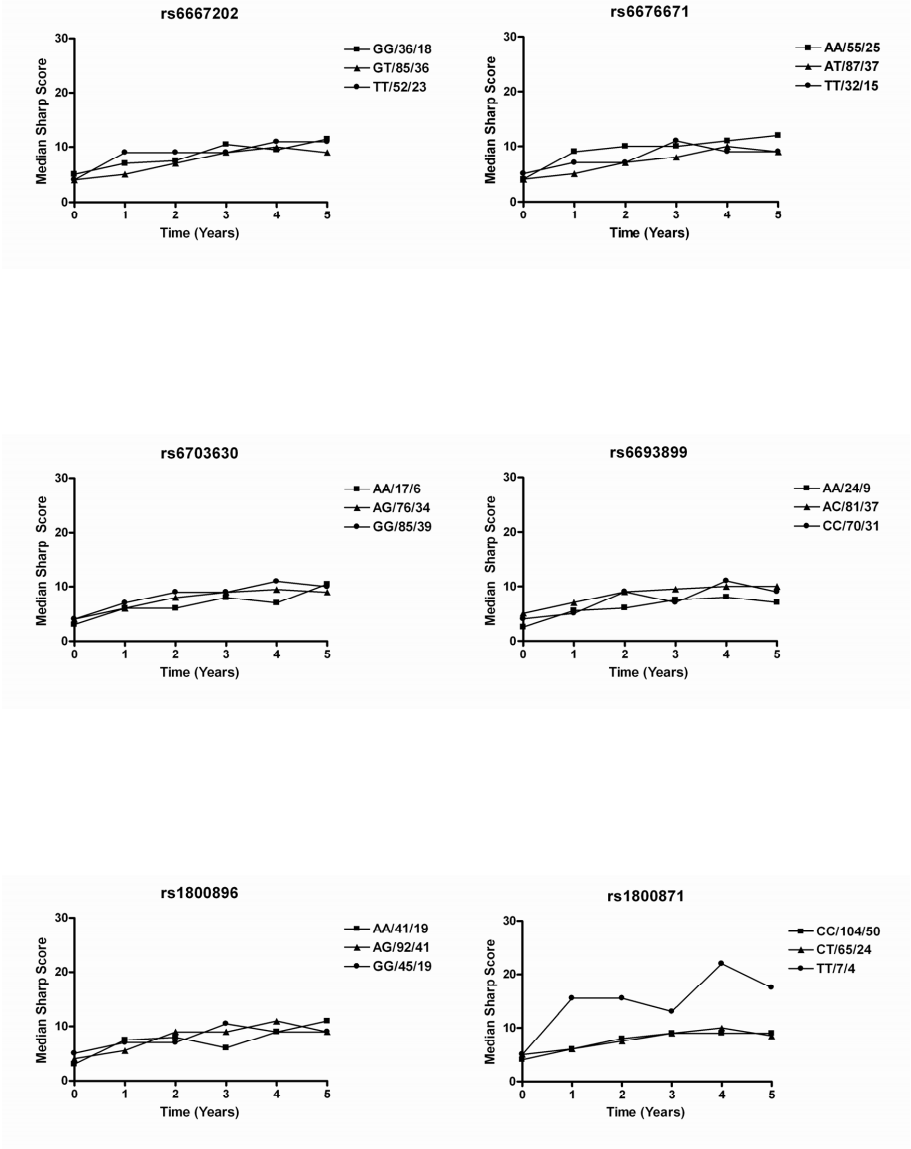
Overview of the raw Sharp-van der Heijde scores, expressed as medians, of all 6 SNPs per genotype for the total patient population. The number of individuals at inclusion (year=0) and after five years of follow-up (years=5) is given in the legend. For example the median sharp van der heijde scores originated from 102 GG, 233GT and 119 TT patients at inclusion and declined to 52 GG, 118 GT and 83 TT after 5 years of follow-up.

Figure 3. Median Sharp-van der Heijde Scores for six IL10 tagging SNPs per genotype in ACPA-positive RA patients



Overview of the raw Sharp-van der Heijde scores, expressed as medians, of all 6 SNPs per genotype for ACPA positive RA patients. The number of individuals at inclusion (year=0) and after five years of follow-up (years=5) is given in the legend. For example the median sharp van der heijde scores originated from 46 GG, 101GT and 73 TT patients at inclusion and declined to 30 GG, 72 GT and 51 TT after 5 years of follow-up.

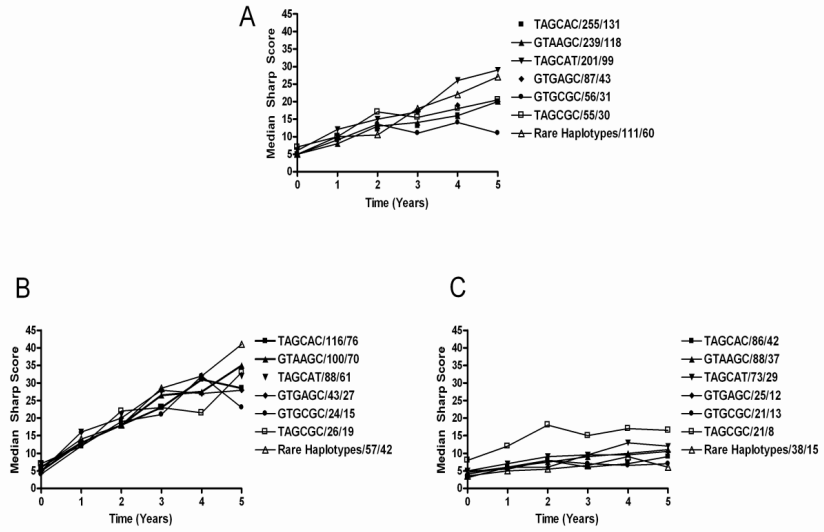
Figure 4. Median Sharp-van der Heijde Scores for six IL10 tagging SNPs per genotype in ACPA-negative RA patients



Overview of the raw Sharp-van der Heijde scores, expressed as medians, of all 6 SNPs per genotype for ACPA negative RA patients. The number of individuals at inclusion (year=0) and after five years of follow-up (years=5) is given in the legend. For example the median sharp van der heijde scores originated from 36 GG, 85 GT and 52 TT patients at inclusion and declined to 18 GG, 36 GT and 23 TT after 5 years of follow-up.

Haplotypes were inferred for 586 individuals from whom genotype data was available in >50% of the SNPs. Haplotypes with a frequency lower than or equal to 2% were pooled into one category termed the rare haplotypes. Six common haplotypes (frequency >2%) capturing 88% of all haplotypes were detected in the total RA population. None of the haplotypes showed a correlation with joint damage in RA patients (Figure 5A). Stratifying RA patients according to ACPA status also revealed no significant correlation (Figure 5 B and C).

Figure 5. Median Sharp-van der Heijde Scores for IL10 haplotypes in (A) all patients (B) ACPA-positive patients and (C) ACPA-negative patients.



Overview of the raw Sharp-van der Heijde scores, expressed as medians, of all 6 SNPs haplotypes for (A) all RA patients (B) ACPA-positive RA patients and (C) ACPA-negative RA patients. The number of haplotypes at inclusion (year=0) and after five years of follow-up (years=5) is given in the legend.

Association of tagging SNPs and haplotypes with remission in RA

Sustained DMARD-free remission was achieved by 13.3% of patients (N=79). Despite a trend towards association between some of the SNPs and sustained DMARD-free remission, none of these associations reached statistical significance (Table 2). Stratification of the RA patients by ACPA status also did not yield any statistically significant associations (Tables 3 and 4). Although there was a possible trend for association between the *IL10* haplotypes and ACPA-negative RA (Table 5), this association was not significant after correction for multiple testing. However, our data is largely underpowered to detect modest effects in this dataset.

Table 2. Association of IL10 SNPs with remission in RA patients ‡.

Genotype		*Remission n=79*	*No remission n=515*	*Hazard Ratio*	*95% CI*	*P*
rs1800871	CC	50 (71%)	269 (60%)			
	CT	18 (26%	153 (34%)			
	TT	2 (3%)	26 (6%)	0.66	0.42-1.04	0.075
rs1800896	AA	13 (18%)	119 (26%)			
	AG	37 (51%)	223 (48%)			
	GG	22 (31%)	123 (27%)	1.20	0.91-1.57	0.28
rs6693899	AA	11 (16%)	62 (14%)			
	AC	33 (47%)	201 (45%)			
	CC	26 (37%)	180 (41%)	0.92	0.70-1.22	0.64
rs6703530*	AA	8 (11%)	45 (9.7%)			
	AG	35 (49%)	177 (38%)			
	GG	29 (40%)	241 (52%)	0.62	0.41-0.94	0.059
rs6676671**	AA	18 (26%)	155 (36%)			
	AT	38 (54%)	193 (44%)			
	TT	14 (20%)	89 (20%)			
rs6667202	GG	18 (27%)	82 (19%)			
	GT	34 (51%)	209 (48%)			
	TT	15 (22%)	145 (33%)	0.73	0.55-0.97	0.064

‡ The number and the percentage of patients are listed, relative to the total number of patients for whom information about the SNP under investigation was available. * Proportional hazards assumption not fulfilled because the small number of patients with genotype AA have an aberrant survival curve. HR therefore only applies to individuals with genotypes AG and GG. ** Proportional hazards assumption not fulfilled, therefore no HR was calculated.

Table 3. Association of IL10 SNPs with remission in ACPA-positive RA patients ‡.

Genotype		*Remission n=9*	*No remission n=264*	*Hazard Ratio*	*95% CI*	*P*
rs1800871	CC	7 (87%)	152 (62%)			
	CT	1 (13%)	80 (32%)			
	TT	0	15 (6.1%)	0.50	0.03-1.84	0.172
rs1800896	AA	1 (11%)	60 (24%)			
	AG	6 (67%)	122 (48%)			
	GG	2 (22%)	70 (28%)	1.11	0.52-2.40	0.82
rs6693899	AA	2 (25%)	34 (14%)			
	AC	4 (50%)	112 (46%)			
	CC	2 (25%)	98 (40%)	0.61	0.27-1.37	0.31
rs6703530*	AA	1 (11%)	25 (9.9%)			
	AG	5 (56%)	96 (38%)			
	GG	3 (33%)	131 (52%)	0.46	0.14-1.53	0.29
rs6676671**	AA	2 (25%)	84 (34%)			
	AT	4 (50%)	110 (45%)			
	TT	2 (25%)	51 (21%)			
rs6667202	GG	2 (22%)	46 (19%)			
	GT	4 (44%)	117 (49%)			
	TT	3 (33%)	77 (32%)	0.99	0.46-2.14	0.98

‡ The number and the percentage of patients are listed, relative to the total number of patients for whom information about the SNP under investigation was available.* Proportional hazards assumption not fulfilled because the small number of patients with genotype AA have an aberrant survival curve. HR therefore only applies to individuals with genotypes AG and GG. ** Proportional hazards assumption not fulfilled, therefore no HR was calculated.

Table 4. Association of IL10 SNPs with remission in ACPA-positive RA patients ‡.

Genotype		*Remission n=61*	*No remission n=146*	*Hazard Ratio*	*95% CI*	*P*
rs1800871	CC	40 (72%)	72 (54%)			
	CT	14 (26%)	55 (41%)			
	TT	1 (2%)	7 (5.2%)	0.58	0.33-0.99	0.045
rs1800896	AA	10 (18%)	36 (26%)			
	AG	27 (48%)	71 (51%)			
	GG	19 (34%)	32 (23%)	1.29	0.95-1.77	0.18
rs6693899	AA	9 (16%)	17 (13%)			
	AC	24 (44%)	63 (47%)			
	CC	22 (40%)	53 (40%)	0.97	0.71-1.34	0.89
rs6703530*	AA	7 (13%)	11 (8.1%)			
	AG	26 (46%)	52 (39%)			
	GG	23 (41%)	72 (53%)	0.66	0.41-1.05	0.14
rs6676671**	AA	14 (25%)	47 (37%)			
	AT	29 (53%)	57 (45%)			
	TT	12 (22%)	22 (18%)			
rs6667202	GG	16 (31%)	19 (15%)			
	GT	24 (47%)	64 (52%)			
	TT	11 (22%)	41 (33%)	0.68	0.49-0.94	0.050

‡ The number and the percentage of patients are listed, relative to the total number of patients for whom information about the SNP under investigation was available.* Proportional hazards assumption not fulfilled because the small number of patients with genotype AA have an aberrant survival curve. HR therefore only applies to individuals with genotypes AG and GG. ** Proportional hazards assumption not fulfilled, therefore no HR was calculated.

Table 5. Association of IL10 haplotypes with remission in all RA patients, ACPA-positive as well as ACPA-negative RA patients.

Haplotype	All Patients		ACPA-Positive Patients		ACPA-Negative Patients	
	No Remission	Remission	No Remission	Remission	No Remission	Remission
TAGCAC	224 (88&)	31 (12%)	122 (96%)	5 (4%)	64 (73%)	24 (27%)
GTAAGC	193 (83%)	41 (17%)	101 (94%)	6 (6%)	58 (65%)	31 (35%)
TAGCAT	187 (90)%	20 (10%)	99 (99%)	1 (1%)	62 (82%)	14 (18%)
GTGAGC	80 (90%)	7 (8%)	45 (98%)	1 (2%)	23 (82%)	5 (18%)
GTGCGC	44 (77%)	13 (23%)	30 (97%)	1 (3%)	9 (43%)	12 (57%)
TAGCGC	52 (85%)	9 (15%)	27 (96%)	1 (4%)	18 (72%)	7 (28%)
Rate (<2%)	100 (84%)	19 (16%)	66 (98.5%)	1 (1.5%)	24 (58.5%)	17 (41.5%)
P (LogRank Test)*		0.07		0.59		0.02

*The log Rank Test was performed since the proportional Hazards Assumption was not met. The log Rank Test assesses whether there is a significant difference in the survival curves estimated by the time to remission.

Discussion

IL10 is an important immunoregulatory cytokine with diverse effects on the immune system. The cytokine is produced by a range of immune cells, including B cells and regulatory T cells(19). Stimulation of human blood cultures with bacterial lipopolysaccharide (LPS) showed large interindividual variation in IL10 secretion, which has been shown to have a genetic component of 50-70%(13). Given the fact that IL10 is produced by a variety of different cells and recent evidence which suggests that the correlation of *IL10 cis* variants with IL10 protein levels can be cell and stimuli specific, it is highly plausible that its regulation is both at the *cis* (caused by variations with the gene region itself) and the *trans* (caused by other variations at other gene loci) level.

Administration of IL10 to animals with collagen-induced arthritis reduces several parameters of the disease and would therefore suggest that high levels of IL10 are likely to result in less progressive disease(30). However, the precise role of IL10 is not clear, even in mouse models. One group recently highlighted this complexity. Hansson and colleagues showed that IL10-deficient mice immunized with collagen type II (CII) develop a more severe disease than their heterozygous littermates(31).

Previous findings in human genetic studies show that the presence of the G allele at *IL10*-2849 (rs6703630) is correlated with enhanced IL10 protein levels in healthy individuals as well as increased levels of autoantibodies and more severe joint damage in RA patients(9). Marinou and colleagues have also recently suggested the involvement of *IL10* polymorphisms in severity of RA(10). The authors suggest that in the absence of autoantibodies (RF or ACPA), patients with a *IL10*-592 (R^2=1 with rs1800871) CC genotype have more severe joint damage in a Polish

population as measured by Larsen's score. However, both studies were conducted in a cross-sectional fashion with data available at one specific time point. The major strengths of the present study are the use of radiological data during long-term follow-up of five years. However, inherent to the design of an inception cohort, not all patients had achieved five years of follow-up, so the number of missing data increased with longer follow-up.

Our data suggests that six *IL10* tagging polymorphisms do not influence radiographic joint damage in RA patients irrespective of the presence or absence of autoantibodies. In addition, haplotypes at this locus also do not show any correlation with severity of disease indicating that a large effect of known *IL10* polymorphisms on joint destruction is unlikely. We have also investigated the role of the *IL10* locus in relation to patients who achieved sustained DMARD-free remission. Our data suggest that individuals with specific *IL10* genotypes or haplotypes do not have a higher rate of remission, irrespective of their ACPA status. However, we cannot exclude that this locus may have a more modest effect on this phenotype. Large well-characterised RA cohorts will be required to investigate the role of modest effect loci in relation to specific phenotypes in RA.

References

(1) Firestein GS. Evolving concepts of rheumatoid arthritis. Nature 2003; 423(6937):356-61.

(2) MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K et al. Characterizing the quantitative

(3) genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum 2000; 43(1):30-7.

(4) Seldin MF, Amos CI, Ward R, Gregersen PK. The genetics revolution and the assault on rheumatoid arthritis. Arthritis Rheum 1999; 42(6):1071-9.

(5) van der Helm-van Mil AH, Kern M, Gregersen PK, Huizinga TW. Variation in radiologic joint destruction in rheumatoid arthritis differs between monozygotic and dizygotic twins and pairs of unrelated patients. Arthritis Rheum 2006; 54(6):2028-30.

(6) Bowes J, Barton A. Recent advances in the genetics of RA susceptibility. Rheumatology (Oxford) 2008; 47(4):399-402.

(7) Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 2008; 40(10):1216-23.

(8) Huizinga TW, Amos CI, van der Helm-van Mil AH, Chen W, van Gaalen FA, Jawaheer D et al. Refining the complex rheumatoid arthritis phenotype based on specificity of the HLA-DRB1 shared epitope for antibodies to citrullinated proteins. Arthritis Rheum 2005; 52(11):3433-8.

(9) 8.Lard LR, van Gaalen FA, Schonkeren JJ, Pieterman EJ, Stoeken G, Vos K et al. Association of the -2849 interleukin-10 promoter polymorphism with autoantibody production and joint destruction in rheumatoid arthritis. Arthritis Rheum 2003; 48(7):1841-8.

(10) Marinou I, Healy J, Mewar D, Moore DJ, Dickson MC, Binks MH et al. Association of interleukin-6 and interleukin-10 genotypes with radiographic damage in rheumatoid arthritis is dependent on autoantibody status. Arthritis Rheum 2007; 56(8):2549-56.

(11) Eskdale J, McNicholl J, Wordsworth P, Jonas B, Huizinga T, Field M et al. Interleukin-10 microsatellite polymorphisms and IL-10 locus alleles in rheumatoid arthritis susceptibility. Lancet 1998; 352(9136):1282-3.

(12) Katsikis PD, Chu CQ, Brennan FM, Maini RN, Feldmann M. Immunoregulatory role of interleukin 10 in rheumatoid arthritis. J Exp Med 1994; 179(5):1517-27.

(13) Westendorp RG, Langermans JA, Huizinga TW, Verweij CL, Sturk A. Genetic influence on cytokine production in meningococcal disease. Lancet 1997; 349(9069):1912-3.

(14) Reuss E, Fimmers R, Kruger A, Becker C, Rittner C, Hohler T. Differential regulation of interleukin-10 production by genetic and environmental factors--a twin study. Genes Immun 2002; 3(7):407-13.

(15) Eskdale J, Keijsers V, Huizinga T, Gallagher G. Microsatellite alleles and single nucleotide polymorphisms (SNP) combine to form four major haplotype families at the human interleukin-10 (IL-10) locus. Genes Immun 1999; 1(2):151-5.

(16) Kurreeman FA, Schonkeren JJ, Heijmans BT, Toes RE, Huizinga TW. Transcription of the IL10 gene reveals allele-specific regulation at the mRNA level. Hum Mol Genet 2004; 13(16):1755-62.

(17) Eskdale J, Gallagher G, Verweij CL, Keijsers V, Westendorp RG, Huizinga TW. Interleukin 10 secretion in relation to human IL-10 locus haplotypes. Proc Natl Acad Sci U S A 1998; 95(16):9465-70.

(18) Forte GI, Scola L, Bellavia D, Vaccarino L, Sanacore M, Sisino G et al. Characterization of two alternative Interleukin(IL)-10 5'UTR mRNA sequences, induced by lipopolysaccharide (LPS) stimulation of peripheral blood mononuclear cells. Mol Immunol 2009.

(19) Mosser DM, Zhang X. Interleukin-10: new perspectives on an old cytokine. Immunol Rev 2008; 226:20518.

(20) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447(7145):661-78.

(21) van Aken J, van Bilsen JHM, Allaart CF, Huizinga TWJ, Breedveld FC. The Leiden Early Arthritis Clinic. Clinical and Experimental Rheumatology 2003; 21(5):S100-S105.

(22) van Aken J, van Bilsen JH, Allaart CF, Huizinga TW, Breedveld FC. The Leiden Early Arthritis Clinic. Clin Exp Rheumatol 2003; 21(5 Suppl 31):S100-S105.

(23) Lard LR, Boers M, Verhoeven A, Vos K, Visser H, Hazes JM et al. Early and aggressive treatment of rheumatoid arthritis patients affects the association of HLA class II antigens with progression of joint damage. Arthritis Rheum 2002; 46(4):899-905.

(24) Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005; 21(2):263-5.

(25) Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007; 81(3):559-75.

(26) van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol 2000; 27(1):261-3.

(27) Bruynesteyn K, van der HD, Boers M, van der LS, Lassere M, van d, V. The Sharp/van der Heijde method out-performed the Larsen/Scott method on the individual patient level in assessing radiographs in early rheumatoid arthritis. J Clin Epidemiol 2004; 57(5):502-12.

(28) van der Heijde D, Boonen A, Boers M, Kostense P, van der Linden S. Reading radiographs in chronological order, in pairs or as single films has important implications for the discriminative power of rheumatoid arthritis clinical trials. Rheumatology (Oxford) 1999; 38(12):1213-20.

(29) Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B et al. The structure of haplotype blocks in the human genome. Science 2002; 296(5576):2225-9.

(30) de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet 2005; 37(11):1217-23.

(31) Trachsel E, Bootz F, Silacci M, Kaspar M, Kosmehl H, Neri D. Antibody-mediated delivery of IL-10 inhibits the progression of established collagen-induced arthritis. Arthritis Res Ther 2007; 9(1):R9.

(32) Hansson AS, Johansson AC, Holmdahl R. Critical role of the major histocompatibility complex and IL-10 in matrilin-1-induced relapsing polychondritis in mice. Arthritis Res Ther 2004; 6(5):R484-R491.

Part III

Other candidate genes in Rheumatoid and Juvenile Arthritis

Chapter 10

Association of tumor necrosis factorα polymorphism and radiographic progression in rheumatoid arthritis: comment on the article by Khanna *et al*.

A.H.M. van der Helm-van Mil
F.A.S. Kurreeman
R.E.M. Toes
T.W.J. Huizinga

Leiden University Medical Center, Leiden, The Netherlands

Arthritis Rheum. ***56****, 1032-1033 (2007)*

To the Editor:

We read with interest the article by Khanna et al regarding the association between the tumor necrosis factor (*TNFA*) -308 polymorphism as well as the HLA–DRB1 shared epitope (SE) alleles and the level of radiographic damage in rheumatoid arthritis (RA) (1). The authors concluded that there is an association between the *TNFA* -308 A allele and the rate of radiographic joint destruction, but that the SE alleles are not associated with the level of joint damage in seropositive RA. In our opinion, both conclusions are questionable. First, Khanna et al investigated the Sharp scores of 189 rheumatoid factor (RF)–positive patients with early RA during 5 years of followup (completed by 45 patients) and compared the scores between patients with and those without SE alleles. The results depended on the method of analysis and on the subgroup of patients. In an analysis unweighted for the number of observations per patient, the presence of SE alleles was not associated with RA severity; in a weighted analysis assessing the total group of patients, the presence of SE alleles was associated with less severe disease; and in an analysis assessing the subgroup of white patients, the presence of SE alleles was associated with more severe joint destruction.

The authors provided no explanation for these discordant findings. The most likely clarification is the fact that the authors selected a group of RF-positive patients. In RA, the presence of RF is highly correlated with the presence of anti–cyclic citrullinated peptide (anti-CCP) antibodies: only 13–18% of RF-positive RA patients are anti-CCP negative (2,3). Furthermore, we recently demonstrated that the SE alleles are primarily a risk factor for the presence of anti-CCP antibodies, and that the presence of SE alleles in RA patients with anti-CCP antibodies is not associated with the development of RA (4). In addition, our analysis on the association between the SE alleles in the presence of anti-CCP antibodies and the Sharp/van der Heijde scores during 4 years of followup was inconclusive; although a trend for more severe joint destruction among SE-positive, anti-CCP–positive patients compared with SE-negative, anti-CCP–positive patients was observed, the number of patients was too low to reach a definite conclusion (5).

The anti-CCP antibody status of patients in the study by Khanna et al is not known, but considering the fact that all patients were RF positive, it is likely that the majority of the patients were anti-CCP positive. The fluctuating results on the absence or presence of an association between SE alleles and the rate of joint destruction in the study by Khanna and colleagues indicates that the sample size of that study was too small to conclude definitely on the association between the SE alleles and the severity of autoantibody-positive RA.

Second, Khanna et al examined whether the *TNFA*-308 G-to-A polymorphism was associated with radiographic joint damage in RF-positive RA. The total Sharp score and the frequency of ≥2 erosions at baseline and during followup were not differently distributed between *TNFA* A and non-A carriers (1). However, when the rate of radiographic progression was derived from the slope of the regression line and weighted for the number of radiographic observations per patient, patients with genotype AA+AG had significantly higher Sharp scores than did patients with genotype GG (1).

The authors also reported that the *TNFA*-308 A alleles are in strong linkage disequilibrium with HLA–DRB1*0301. Because it was recently reported in North American and Dutch patients that the HLA–DRB1*0301 allele (and the A1;B8;DR3 haplotype)is associated with anti-CCP–negative RA (2,3), and, because anti-CCP–negative RA is associated with less severe joint damage, the observation that the *TNFA* -308 A allele is associated with a higher rate of joint destruction seems contrasting. Considering the linkage disequilibrium between the *TNFA* -308 A allele and HLA–DRB1*0301, the analysis on the association between this polymorphism and the rate of joint destruction should be corrected for the presence or absence of anti-CCP antibodies.

Therefore, we determined the *TNFA* -308 polymorphism in 327 white patients who presented to the Leiden Early Arthritis Clinic between 1993 and 2000 and in whom RA was diagnosed during the first year after inclusion. Eight patients (2%) had genotype AA, 94 patients (29%) had genotype AG, and 225 patients (69%) had genotype GG. Anti-CCP antibodies were present in 182 patients (56%). Radiographs were available for 267 patients after 1 year of followup, for 205 patients after 2 years, and for 154 patients after 4 years of followup. The Sharp/van der Heijde scores were not significantly different between the RA patients who carried a *TNFA* -308 A allele and the patients who did not carry an A allele (Figure 1). Subsequently, this analysis was performed in subgroups of patients with and those without anti-CCP antibodies; the results revealed no significant differences. A linear regression analysis with the radiographic progression score as dependent variable and age, sex, *TNFA* _308 A/non-A carriership, and anti-CCP antibody status as possible explanatory variables revealed that only the presence of anti-CCP antibodies (regression coefficient 9.0, standard error 2.3, P <0.001) and age (regression coefficient 0.2, standard error 0.07, P < 0.003) were significantly associated with radiographic joint destruction.

In our opinion, because the *TNFA* -308 A allele is in linkage disequilibrium with HLA–DR3, which is reported to be associated with only anti-CCP–negative RA, and RA patients without anti-CCP antibodies generally experience a less severe disease course compared with RA patients who have anti-CCP antibodies, the analysis of the association between the *TNFA* -308 A allele and the rate of joint destruction should be corrected for the presence or absence of anti-CCP antibodies. Therefore, we analyzed white RA patients who were included in the Leiden Early Arthritis Clinic, and we did not observe an association between the *TNFA* -308 A allele and radiographic joint destruction in RA patients, neither in the presence nor in the absence of anti-CCP antibodies.

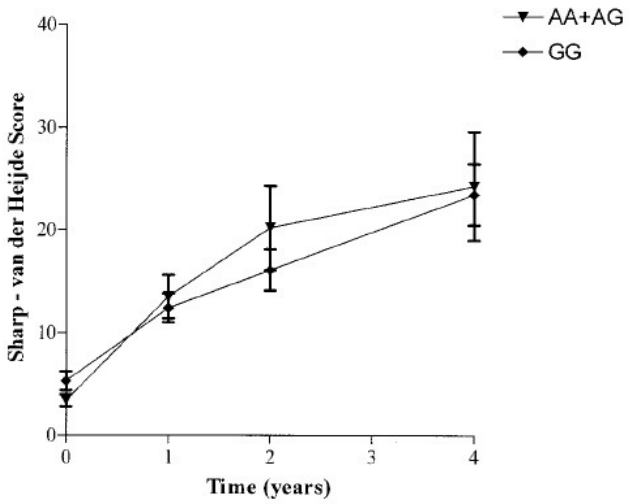


Figure 1. Sharp/van der Heijde scores (mean ± SEM) in rheumatoid arthritis patients with *TNFA* −308 genotype AA + AG and those with genotype GG, during 4 years of followup. P = 0.8, 0.7, 0.8, and 0.7, respectively, at baseline and at 1 year, 2 years, and 4 years of followup, by Mann-Whitney test.

Chapter 11

Association of the 6q23 region with the rate of joint destruction in rheumatoid arthritis

H.U. Scherer
M.P.M. van der Linden
F.A.S. Kurreeman
G. Stoeken-Rijsbergen
S. le Cessie
T.W.J. Huizinga
A.H.M. van der Helm-van Mil
R.E.M. Toes

Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands
Department of Medical Statistics and Bioinformatics, Leiden University Medical Centre, Leiden, The Netherlands
Department of Rheumatology and Clinical Immunology, Charité University Medicine Berlin, Berlin, Germany

Ann Rheum Dis. (2009)

Abstract

Background

Two novel genetic polymorphisms on chromosome 6q23 are associated with susceptibility to rheumatoid arthritis (RA). Both polymorphisms (rs6920220 and rs10499194) reside in a region close to the gene encoding tumor necrosis factor α-induced protein 3 (TNFAIP3). TNFAIP3 is a negative regulator of NFκB and as such involved in inhibiting TNF-Receptor mediated signalling effects. Interestingly, the initial associations were detected in patients with long-standing RA. However, no association was found for rs10499194 in a Swedish early arthritis cohort. As this could be caused by overrepresentation of patients with severe disease in cohorts with long-standing RA, we analyzed the effect of the 6q23 region on the rate of joint destruction.

Methods and Findings

Five single nucleotide polymorphisms (SNPs) in 6q23 were genotyped in 324 Dutch patients with early RA. Genotypes were correlated to progression of radiographic joint damage for a follow-up time of 5 years. Two polymorphisms (rs675520 and rs9376293) associated with severity of radiographic joint damage in ACPA+ patients. Importantly, the effects were present after correction for confounding factors such as secular trends in treatment

Conclusions

Our data associate the 6q23 region with the rate of joint destruction in ACPA+RA.

Introduction

Recent whole genome association scans have revealed novel genetic polymorphisms associated with susceptibility to ACPA+ RA.[1,2] Among those, two single nucleotide polymorphisms (SNPs), rs6920220 (A allele) and rs10499194 (C allele), were found to independently associate with ACPA+ disease. Both SNPs map to a single linkage disequilibrium block spanning ~60 kb in a region on chromosome 6q23 that lacks known genes or transcripts. The closest genes are oligodendrocyte lineage transcription factor 3 (OLIG3) and tumor necrosis factor α-induced protein 3 (TNFAIP3). The latter is of potential importance to RA pathogenesis, as the protein TNFAIP3 acts as a negative regulator of NF-κB.[3] So far, however, functional relevance of the reported polymorphisms is unknown.

Rs6920220 was initially identified in ACPA+ RA patients (minor allele OR 1.38) originating from the United Kingdom (UK).[1] It was further replicated in an extended UK based case-control study.[4] Rs10499194 was initially identified in North American ACPA+ patients (the Brigham Rheumatoid Arthritis Sequential Study, BRASS; minor allele OR 0.67).[2] Replication was successful in two additional US cohorts selected from the North American Rheumatoid Arthritis Consortium (NARAC). Replication failed, however, in ACPA+ patients of a Swedish population-based inception cohort (the Epidemiological Investigation of Rheumatoid Arthritis cohort, EIRA).[2] This latter finding is of interest, as both BRASS and NARAC are cohorts of patients with long-standing RA (mean disease duration BRASS: 15.4 ± 12.8 years [5]; NARAC: 14.3 ± 11.1 years [6]). The EIRA study, however, was designed to identify incident cases of RA as soon as possible after disease onset, resulting in an estimated mean disease duration at inclusion of only 10 months.[7]

Association of a genetic polymorphism in cohorts of patients with longstanding disease but absence of this association in an early arthritis cohort led us to hypothesize that the 6q23 region would associate with disease severity in ACPA+ patients. Very little information is currently available on the effects of genetic variation on outcome measures in RA [8]. Therefore, we genotyped five SNPs in a Dutch early arthritis cohort (the Leiden Early Arthritis Clinic, EAC) and correlated genotyping data to progression of radiographic joint damage for a maximum follow up of 5 years.

Patients & Methods

Patients

The Leiden EAC is a population-based inception cohort that includes patients with self-reported symptom duration of ≤ 2 years.[9] DNA samples of 324 patients consecutively included between 1993 and 2003 were used for analysis. For further details see supplementary file 1.

SNP selection and genotyping

Five SNPs (rs1878658, rs675520, rs9376293, rs10499194 and rs6920220) were selected based on a haplotype analysis across the 6q23 locus published previously.[2] All SNPs are in imperfect linkage disequilibrium to one another (supplementary table 1). Genotyping was performed using pre-designed Taqman allelic discrimination probes (Applied Biosystems).

Each 384 well plate contained 10 ng sample DNA per well and at least 8 negative and 6 positive controls. Genotype calls and clusters were manually checked for discrepancies and doubtful calls were rejected. No SNP deviated from Hardy-Weinberg equilibrium. Genotyping call rates were 96.5 % (rs1878658), 98 % (rs675520), 95 % (rs9376293), 94 % (rs10499194), and 98.1 % (rs6920220).

Serology and Radiographs

Serum samples were tested for citrulline-specific IgG antibodies using a commercially available ELISA kit (Immunoscan Mark2, Eurodiagnostica, The Netherlands). Radiographs were scored according to the Sharp van der Heijde method [10] with known time order by one blinded, independent trained reader (supplementary file 1).

Statistical Analysis

Association between genotypes and radiographic scoring data was analyzed using SPSS version 16.0 (SPSS Inc., Chicago, IL). P-values < 0.05 were considered significant. All p-values reported are two-sided.

Two approaches were chosen for statistical analysis. First, the average increase in Sharp van der Heijde scores during the follow-up period was estimated per person by regression analysis. Subsequently, the average increase (slope) of scores per genotype was compared non-parametrically using the Mann-Whitney rank-sum test.

We observed an influence of the time of inclusion (1993-2003) on the progression of radiographic joint damage reflecting most likely an improvement of treatment intensity during this 10 year time period. In order to account for this effect, we performed, as a second approach, a mixed model analysis described in detail in supplementary file 1.

Results

Radiographic scores of 324 Dutch RA patients (181 ACPA+, 143 ACPA-) were available for analysis. At least five radiographic follow-up observations were available in 57% of patients. A dominant model was chosen for analysis, as the frequency of patients homozygous for the minor allele of rs1878658 (G), rs10499194 (T) and rs6920220 (A) was ≤ 5%. Figure 1 depicts the influence of genotypes on radiographic joint damage. ACPA+ and ACPA- subgroups were analyzed separately. Median scores and interquartile ranges (IQR) are provided for ACPA+ patients in table 1 (for ACPA- patients see supplementary table 2).

No influence of genotypes on radiographic joint damage was observed in ACPA- patients (figure 1). In ACPA+ patients, however, two polymorphisms showed reproducible association with disease progression over time. Presence of the G allele of rs675520 was found to associate with increased Sharp van der Heijde scores, as a significant difference was observed when the average increase (slope) in radiographic scores over time was compared with G as the dominant allele (median slope AG/GG = 4.6, AA = 2.3; Mann-Whitney p = 0.007). In order to account for an effect of improving treatment strategies on radiographic progression during the 10 year period in which patients were included into the study, we next performed a mixed model analysis. This analysis identified the year of inclusion as a significant variable influencing the extent of radiographic joint damage (p = 0.005). After correcting for the year of inclusion, however, we still observed a significant influence of the G allele of rs675520 (AG/GG vs. AA, p = 0.026).

Similar to the G allele of rs675520, we noted an influence of the C allele of rs9376293 on progression of radiographic joint damage (figure 1). The average increase (slope) in Sharp van der Heijde scores over time was significantly higher for C allele carriers as compared to T homozygotes (median slope CC/CT = 4.5, median slope TT = 3.0, Mann-Whitney p = 0.021). After correcting for the year of inclusion as described above a trend effect of the C allele remained (p = 0.097). For rs1878658, rs10499194 and rs6920220, no significant influence of individual genotypes on radiographic joint damage was noted.

Figure 1

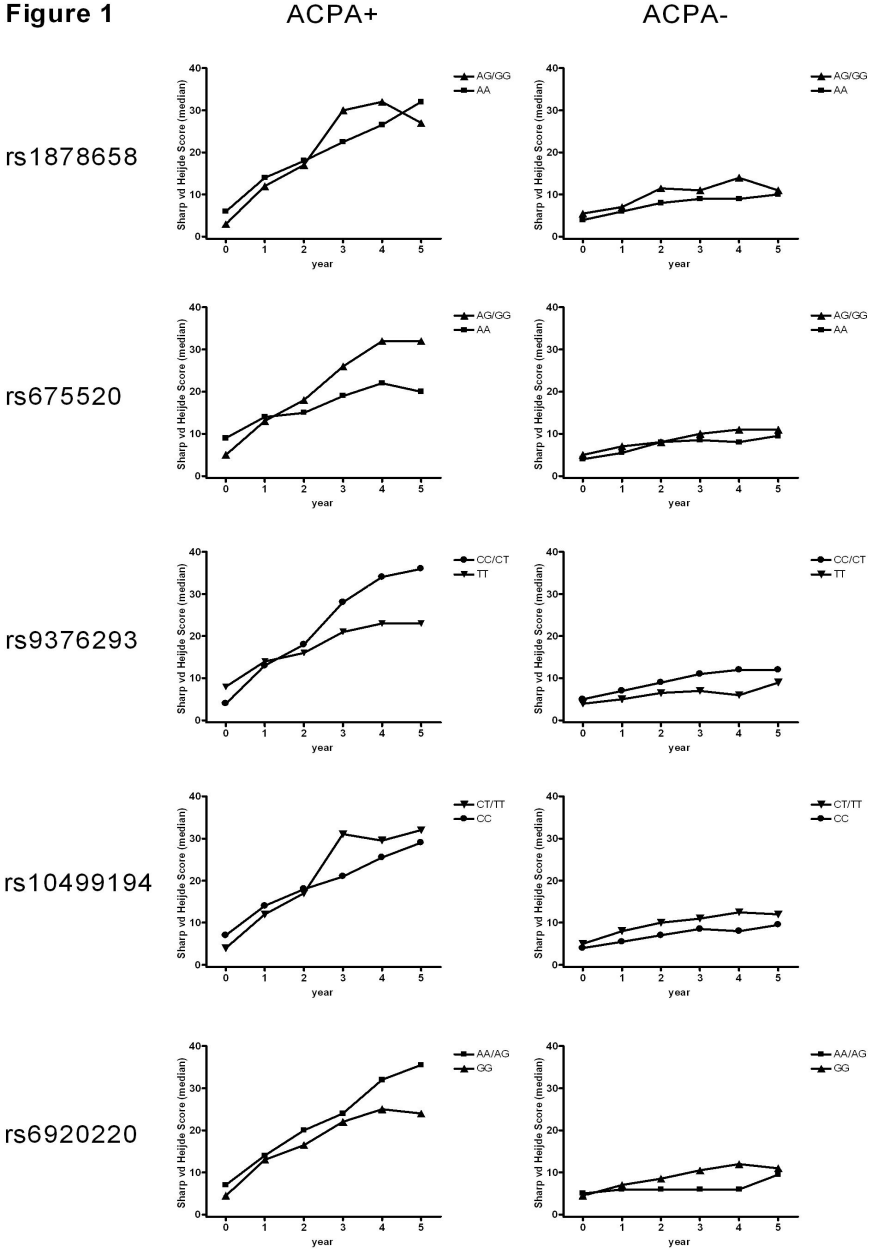


Figure 1. Development of median Sharp van der Heijde scores plotted according to genotype/allele in ACPA+ (left column) and ACPA- (right column) RA patients. Year 0 equals baseline-values. Regression analysis was performed in order to estimate the average increase (slope) in Sharp van der Heijde scores over time. Slopes were subsequently compared using the nonparametric Mann-Whitney test (for the ACPA+ subgroup: p = 0.37 (rs1878658); p = 0.007(rs675520); p = 0.021 (rs9376293); p = 0.05 (rs10499194); p = 0.76 (rs6920220)).

Table 1: Median Sharp van der Heijde scores (M) and interquartile ranges (IQR; 25 - 75% percentiles) per genotype for ACPA+ RA patients (# = number of patients). Genotypes were combined for rs1878658 (G), rs10499194 (T) and rs6920220 (A), as the frequency of patients homozygous for the respective minor allele was ≤5% for these SNPs.

	year																	
	0			1			2			3			4			5		
ACPA+	**M**	**IQR**	**#**	**M**	**IQR**	**#**	**M**	**IQR**	**#**	**M**	**IQR**	**#**	**M**	**IQR**	**#**	**M**	**IQR**	**#**
rs1878658																		
AA	6	2 - 13	126	14	6 - 24	111	18	10 - 32	113	22.5	11 - 42	98	26.5	14 - 51.5	86	32	14 - 51.5	89
AG/GG	3	1 - 7	40	12	5.5 - 20.5	40	17	6 - 32	39	30	11.5 - 56.5	34	32	16 - 69	31	27	17 - 74	31
rs675520																		
AA	9	4.5 - 13	33	14	6.5 - 21	32	15	8 - 27.5	33	19	9 - 34	27	22	13.5 - 39.5	24	20	12.5 - 47.5	21
AG	5	1 - 10	83	12	5 - 21	73	18	9.5 - 32.5	78	25	11.5 - 54	68	32.5	14 - 61	58	33	17 - 62.5	66
GG	5	2 - 11	51	14	7 - 27	48	18	8 - 37	43	26	14 - 48	39	28	16 - 63	37	32	20 - 64	35
AG/GG	5	2 - 10	134	13	6 - 24.5	121	18	9.5 - 33	121	26	12 - 53	107	32	15 - 61	95	32	18.5 - 61	101
rs9376293																		
CC	6.5	3 - 13.5	28	15	9 - 28	27	27	12 - 49.5	25	33	15.5 - 58	24	35	16.5 - 62	24	35	16.5 - 56	21
CT	4	1 - 9	84	12	5 - 21	79	17	7 - 32.5	78	21	10.5 - 54	66	32	14 - 68	57	36	20 - 68	61
TT	8	3 - 13	54	14	6 - 21	46	16	10.5 - 25.5	50	21	12 - 30.5	44	23	13.5 - 38	38	23	13 - 47	40
CC/CT	4	2 - 10	112	13	6 - 26.5	106	18	8 - 36	103	28	11 - 54.5	90	34	15 - 63	81	36	19 - 67.5	82
rs10499194																		
CC	7	2 - 12.5	90	14	6 - 21.5	77	18	9.5 - 28	82	21	12 - 35	71	25.5	14 - 45	62	29	14 - 48	62
CT/TT	4	2 - 7.5	65	12	6 - 24.5	64	17	7 - 37	61	31	11 - 55	55	29.5	14.5 - 67	50	32	15 - 68	51
rs6920220																		
AA/AG	7	2 - 12	58	14	6 - 21	53	20	10 - 32.5	57	24	12 - 48	47	32	18 - 56	35	35.5	18 - 65	42
GG	4.5	2 - 10.5	108	13	6 - 25	99	16.5	8 - 30.5	96	22	11 - 48	86	25	14 - 61	84	24	14 - 51	79

Discussion

The 6q23 region has recently been associated with disease susceptibility in RA. This region contains no known transcripts. The closest genes with known function are OLIG3 and TNFAIP3. TNFAIP3 encodes protein A20, a TNF-α induced negative regulator of NF-κB.[3,11] Decreased levels of A20 lead to uncontrolled NFκB-activity, resulting in increased inflammation. This observation makes TNFAIP3/A20 and the 6q23 region interesting candidates that could modulate inflammation also in RA.

We were intrigued by recent differential findings for rs10499194, a SNP on chromosome 6q23 close to TNFAIP3, in cohorts with differing disease duration. The major allele (C) was found to associate with disease susceptibility in ACPA+ RA patients in three cohorts with long-standing disease, but not in an early arthritis cohort.[2] This indicated a potential impact of the 6q23 region on disease severity. In order to test for such an impact, five SNPs were genotyped in a cohort of Dutch patients with early RA. These SNPs had previously been shown to identify common haplotypes in 6q23.[2] We identified two SNPs for which presence of alleles was associated with increased joint destruction in ACPA+ patients. Carriers of the G allele of rs675520 developed increased Sharp van der Heijde scores over time. A similar effect, although weaker, was found for the C allele of rs9376293. Interestingly, no association was found for any of the SNPs in ACPA- individuals. Although this does not exclude a contribution of the 6q23-region to disease severity in ACPA- disease, the latter observation is in line with recent reports detecting an association of the 6q23 region with disease susceptibility in ACPA+ patients only.[4] No effect on disease severity was observed for rs10499194 and rs6920220. Based on our data we cannot rule out the possibility that either SNP exerts a weak effect that requires larger sample numbers for detection or that cannot be observed during the first years of disease. Interestingly, we observed nominally higher scores for the risk conferring A allele of rs6920220 without reaching statistical significance. The discrepancy between SNPs associating with susceptibility and radiographic progression also indicates that the causal variant at this locus has not yet been identified. Given the large area of linkage disequilibrium surrounding these SNPs, further fine-mapping and functional characterization will have to be performed.

Data linking newly identified genetic polymorphisms to disease outcome in RA are only beginning to emerge. Our data are unique, as they cover a long period of radiographic follow-up and have been scrutinized for artefacts such as secular trends in treatment intensity. Albeit based on relatively low patient numbers, our data indicate a contribution of the 6q23 region to the rate of joint destruction in ACPA+ RA, thereby further refining our understanding of the effects exerted by this locus. Replication of our findings in other cohorts is needed. Nonetheless, this is the first study demonstrating such an effect for genetic polymorphisms located outside the HLA-region in ACPA+ RA patients.

References

(1) Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447(7145):661-678.

(2) Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007; 39(12):1477-1482.

(3) Wertz IE, O'Rourke KM, Zhou H, Eby M, Aravind L, Seshagiri S et al. Deubiquitination and ubiquitin ligase domains of A20 downregulate NF-kappaB signalling. Nature 2004; 430(7000):694-699.

(4) Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J et al. Rheumatoid arthritis association at 6q23. Nat Genet 2007; 39(12):1431-1433.

(5) Karlson EW, Chibnik LB, Cui J, Plenge RM, Glass RJ, Maher NE et al. Associations between human leukocyte antigen, PTPN22, CTLA4 genotypes and rheumatoid arthritis phenotypes of autoantibody status, age at diagnosis and erosions in a large cohort study. Ann Rheum Dis 2008; 67(3):358-363.

(6) Jawaheer D, Lum RF, Amos CI, Gregersen PK, Criswell LA. Clustering of disease features within 512 multicase rheumatoid arthritis families. Arthritis Rheum 2004; 50(3):736-741.

(7) Stolt P, Bengtsson C, Nordmark B, Lindblad S, Lundberg I, Klareskog L et al. Quantification of the influence of cigarette smoking on rheumatoid arthritis: results from a population based case-control study, using incident cases. Ann Rheum Dis 2003; 62(9):835-841.

(8) van der Helm-van Mil AH, Padyukov L, Toes RE, Klareskog L, Huizinga TW. Genome-wide single-nucleotide polymorphism studies in rheumatology: Hype or hope? Arthritis Rheum 2008; 58(9):2591-2597.

(9) van Aken J, van Bilsen JH, Allaart CF, Huizinga TW, Breedveld FC. The Leiden Early Arthritis Clinic. Clin Exp Rheumatol 2003; 21(5 Suppl 31):S100-S105.

(10) van der Heijde DM. Plain X-rays in rheumatoid arthritis: overview of scoring methods, their reliability and applicability. Baillieres Clin Rheumatol 1996; 10(3):435-453.

(11) Opipari AW, Jr., Boguski MS, Dixit VM. The A20 cDNA induced by tumor necrosis factor alpha encodes a novel type of zinc finger protein. J Biol Chem 1990; 265(25):14705-14708.

Chapter 12

Confirmation of STAT4, IL2/IL21, and CTLA4 Polymorphisms in Rheumatoid Arthritis

N.A. Daha
F.A.S. Kurreeman
R.B. Marques
G. Stoeken-Rijsbergen
W. Verduijn
T.W.J. Huizinga
R.E.M. Toes

Leiden University Medical Center, Leiden, Netherlands

Arthritis Rheum. 60, 1255-60 (2009)

Abstract

Objective

Recent advances have led to novel identification of genetic polymorphisms that are associated with susceptibility to rheumatoid arthritis (RA). Currently, 5 loci (HLA, PTPN22, TRAF1/C5, TNFAIP3, and STAT4) have been consistently reported, whereas others have been observed less systematically. The aim of the present study was to independently replicate 3 recently described RA susceptibility loci, STAT4, IL2/ IL21, and CTLA4, in a large Dutch case-control cohort, and to perform a meta-analysis of all published studies to date and investigate the relevance of the findings in clinically well-defined subgroups of RA patients with or without autoantibodies

Methods

The STAT4, IL2/IL21, and CTLA4 gene polymorphisms (rs7574865, rs6822844, and rs3087243, respectively) were genotyped in 877 RA patients and 866 healthy individuals. A meta-analysis of all published studies of disease association with these polymorphisms was performed using the Mantel-Haenszel fixed-effects method.

Results

An association of STAT4, IL2/IL21, and CTLA4 with RA was detected in Dutch patients (odds ratio [OR] 1.19 [P = 0.031], OR 0.84 [P = 0.051], and OR 0.87 [P = 0.041], respectively). Results from the meta-analysis confirmed an association of all 3 polymorphisms with RA in Caucasians (OR 1.24 [P = 1.66 x10^{-11}], OR 0.78 [P = 5.6 x 10^{-5}], and OR 0.91[P = 1.8 x 10^{-3}], respectively). The meta-analysis also revealed that STAT4 predisposed to disease development equally in patients with autoantibodies and those without autoantibodies, and that CTLA4 enhanced the development of anti–citrullinated protein antibody (ACPA)-positive RA as compared with ACPA-negative RA.

Conclusion

Our results replicate and firmly establish the association of STAT4 and CTLA4 with RA and provide highly suggestive evidence for IL2/IL21 loci as a risk factor for RA. Given the strong statistical power of our meta-analysis to confirm a true-positive association, these findings provide considerable support for the involvement of CTLA4 in distinct subsets of RA patients.

Introduction

Rheumatoid arthritis (RA) is a common autoimmune disease with unknown etiology. Nonetheless, it is known that both genetic and environmental factors play a role in the pathogenesis of the disease. The strongest known genetic association with RA is with particular alleles of the HLA locus (1). In recent years, continuing advances in genotyping techniques have led to discovery of a large number of potential genetic associations outside this region (2–4). Some of these newly identified susceptibility loci represent true associations, whereas others still remain to be conclusively investigated.

Follow up replication studies in different populations are needed to resolve this issue. However, although some follow up studies have shown robust associations, others have yielded encouraging, but inconsistent, results. This could be due to insufficient power to detect modest effects in some of these studies. To overcome this limitation, data from previously published studies can be systematically evaluated by a meta-analysis. Furthermore, since RA is a heterogeneous disease and data indicate that different risk factors predispose to autoantibody-positive disease as compared with autoantibody-negative disease (5), further investigation in these disease subsets remains to be performed in large data sets. In the present study, 3 previously described susceptibility loci in patients with RA, i.e., rs7574865 (for signal transducer and activator of transcription 4 [*STAT4*]), rs6822844 (for interleukin-2/interleukin-21 [*IL2/IL21*]), and rs3087243 (for cytotoxic T lymphocyte– associated antigen 4 [*CTLA4*]), were investigated for association with the disease and for association with autoantibody status.

The association of *STAT4* with RA was first described in 2007, followed by a vast number of replication studies in both Caucasian and East Asian populations, all of which yielded results that were consistent across the studies (3,6–11). In contrast, the second polymorphism in this study, namely, rs6822844 in the *IL2*/*IL21* region, has only been described in one study thus far, indicating that further replication is needed (4). The third variant that we aimed to investigate maps to the *CTLA4* region. Various replication studies have shown encouraging, yet inconsistent, results for this locus (12–17). Plenge et al provided evidence that the differences between studies could be due to insufficient power in some of the studies (15).

Thus, the aim of our study was to replicate 3 previously described risk factors for RA and further study their association by a meta-analysis. Additionally, we examined whether the association was restricted to clinically relevant disease subsets that were characterized by autoantibody status.

Patients & Methods

Patients

A total of 877 RA patients whose diagnosis met the American College of Rheumatology (formerly, the American Rheumatism Association) 1987 revised criteria for RA (18) were recruited in 1994 from hospitals in the western part of The Netherlands, of whom 602 patients were from 2 independent cohorts of patients with early arthritis, the EAC and BeSt cohorts, and 275 were from the outpatient clinic of the Leiden University Medical Center. For both early arthritis cohorts, anti–citrullinated protein antibody (ACPA) status and rheumatoid factor (RF) status were obtained. For the patients from the outpatient clinic, only the RF status was obtained. Patients' characteristics have been described previously (19). As healthy controls, 866 subjects were randomly selected from the Immunogenetics and Transplantation Immunology section of Leiden University Medical Center. All patients and controls gave their informed consent to participate in the study, and the study was approved by the local ethics committee of the participating hospitals.

Genotyping methods

Genotyping of *STAT4* rs7574865, *IL2*/*IL21* rs6822844, and *CTLA4* rs3087243 was performed using MassArray matrix-assisted laser desorption ionization– time-of-flight mass spectrometry, according to the protocols recommended by the manufacturer (Sequenom, San Diego, CA). SpectroCaller software, which was supplied by the same manufacturer, was used to automatically identify, i.e., call, the genotypes. Each 384-well plate consisted of 8 positive controls and 8 negative controls, all of which were indeed shown to be positive or negative. Clusters were evaluated and all doubtful calls were checked; after manually evaluating the spectra of each cluster, the genotypes were accepted, recalled, or rejected. At least 10% of the genotypes were assessed in duplicate, with an error rate of <1%.

Statistical analysis

Allele distribution was analyzed for association with RA using a chi-square test with 1 degree of freedom. Odds ratios (ORs) and 95% confidence intervals (95% CIs) were calculated using the Statcalc module of Epi Info software (Centers for Disease Control and Prevention, Atlanta, GA). *P* values less than 0.05 were considered significant. Genotype frequencies in cases and controls did not deviate from Hardy-Weinberg equilibrium at a significance level of $P < 0.05$.

A meta-analysis of published reports describing disease associations with *STAT4*, *IL2*/*IL21*, and *CTLA4* was performed using the Mantel-Haenszel method of combining ORs. Reports published up to July 31, 2008 were included in the analysis. Heterogeneity of the ORs across sample sets was analyzed using the Breslow-Day test.

Since no significant heterogeneity was observed among the studies, the ORs and 95% CIs were calculated using a fixed-effects model, and *P* values less than 0.05 were considered significant. (Genotype frequencies are available upon request from the corresponding author.) The meta-analyses had >80% power to detect allele associations both for association with RA and for association with autoantibody status, at ORs of ≤1.18 (at a significance level of $P < 0.05$) for all 3 polymorphisms, except for the association of the *IL2*/*IL21* locus with autoantibody status, in which the meta-analysis had 67% power to detect an OR of 1.20 (at a significance level of $P < 0.05$).

Results

Replication of STAT4, IL2/IL21, and CTLA4 loci in an independent Dutch cohort

The polymorphisms *STAT4* rs7574865, *IL2*/*IL21* rs6822844, and *CTLA4* rs3087243 were genotyped in 877 RA patients and 866 healthy controls. Both *STAT4* and *CTLA4* showed an association with RA in the Dutch cohort, while a clear trend toward association was observed for the *IL2*/*IL21* locus (Table 1). The results had the same direction of association as has been reported in previous studies (3,4,6,9–12,14,15,17).

Stratification by autoantibody status

For a better understanding of disease etiology, we investigated whether the associations were restricted to a specific subset of disease, defined by either ACPA positivity or RF positivity. In these subsets, no differential association could be observed for either the *STAT4* polymorphism or the *IL2*/*IL21* polymorphism (Table 1). The *CTLA4* polymorphism, however, did show a significant association with ACPA-positive RA in patients as compared with healthy controls (OR 0.80, 95% CI 0.66–0.96, $P = 0.015$), but showed no association with ACPA-negative RA in patients as compared with healthy controls (OR 0.90, 95% CI 0.70–1.12, $P = 0.325$). Furthermore, an increase in frequency of the G allele was observed in patients with ACPA-positive RA (60%) as compared with patients with ACPA-negative RA (57%), but the difference was not significant.

Table 1. Results of association and stratification analysis of *STAT4* rs7574865, *IL2*/*IL21* rs6822844, and *CTLA4* rs3087243 in a Dutch cohort*

	STAT4 rs7574865					*IL2*/*IL21* rs6822844					*CTLA4* rs3087243				
	Allele G	Allele T	MAF	OR (95% CI)	*P*	Allele G	Allele T	MAF	OR (95% CI)	*P*	Allele A	Allele G	MAF	OR (95% CI)	*P*
RA															
Cases	1,276	432	0.25	1.19 (1.01–1.40)	0.031	1,469	285	0.16	0.84 (0.70–1.0)	0.0506	729	1,005	0.42	0.87 (0.76–1.00)	0.041
Controls	1,348	384	0.22			1,407	325	0.19			785	941	0.45		
RF status															
RF+	711	239	0.25	1.18 (0.98–1.43)	0.080	811	163	0.17	0.87 (0.70–1.08)	0.188	406	554	0.42	0.88 (1.75–1.03)	0.111
RF−	362	116	0.24	1.12 (0.88–1.44)	0.332	422	78	0.16	0.80 (0.61–1.06)	0.105	207	287	0.42	0.86 (0.70–1.06)	0.159
RF+ vs. RF−				1.05 (0.81–1.37)	0.713				1.09 (0.80–1.48)	0.577				1.02 (0.81–1.27)	0.887
ACPA status															
ACPA+	478	158	0.25	1.16 (0.93–1.44)	0.170	552	102	0.16	0.80 (0.62–1.03)	0.072	257	387	0.40	0.80 (0.66–0.96)	0.015
ACPA−	325	111	0.25	1.20 (0.93–1.54)	0.144	380	76	0.17	0.87 (0.65–1.15)	0.303	193	257	0.43	0.90 (0.70–1.12)	0.325
ACPA+ vs. ACPA−				0.97 (0.72–1.29)	0.819				0.92 (0.66–1.29)	0.633				0.88 (0.69–1.14)	0.324

* Values for alleles G, T, and A are the allele frequency. MAF = minor allele frequency; OR = odds ratio; 95% CI = 95% confidence interval; RA = rheumatoid arthritis; RF = rheumatoid factor; ACPA = anti–citrullinated protein antibody.

Meta-analysis of STAT4, IL2/IL21, and CTLA4 loci

Association with RA overall.

To systematically assess the contribution of the 3 studied polymorphisms in RA, a meta-analysis of all published studies to date was performed. This analysis provided an overall OR for the widely and consistently replicated STAT4 locus in the Caucasian population (OR 1.24, 95% CI 1.17–1.33, P = 1.66 x 10^{-11}) (Table 2). Evaluation of both studies dealing with IL2/IL21 provided additional evidence of an association of IL2/IL21 with RA in Caucasians (OR 0.78, 95% CI 0.69–0.88, P = 5.6 x 10^{-5}). Examination of the 6 previously published studies on rs3087243 in CTLA4 confirmed an overall association of this region with RA in Caucasians (OR 0.91, 95% CI 0.85–0.96, P = 1.8 x 10^{-3}) (Table 2).

Association with RA stratified by autoantibody status.

Results from our meta-analysis indicated that STAT4 was associated with both autoantibody-positive and autoantibody-negative disease in the Caucasian population (OR 1.00, 95% CI 0.89–1.14, P = 0.97) (Table 3). Moreover, IL2/IL21 showed a significant association with RF-positive disease in Caucasian patients as compared with healthy controls (OR 0.78, 95% CI 0.68–0.90, P = 6.9 x 10^{-4}), but IL2/IL21 showed no significant association with RF-negative disease in Caucasian patients as compared with controls (OR 0.82, 95% CI 0.66–1.03, P = 0.083). However, the effect sizes of both associations were of the same extent. Furthermore, the effect size of RF-positive disease compared with RF-negative disease was limited, indicating an association of IL2/IL21 in both disease subsets (OR 1.02, 95% CI 0.81–1.29, P = 0.86) (Table 3). Interestingly, in the meta-analysis, CTLA4 in the Caucasian population was found to predispose to ACPA-positive disease only, and not to ACPA-negative disease (OR 0.86, 95% CI 0.78–0.96, P = 4.7 x 10^{-3}) (Table 3).

Table 2. Meta-analysis of the association of *STAT4* (rs7574865), *IL2/IL21* (rs6822844), and *CTLA4* (rs3087243) with RA in Caucasian and East Asian populations*

	RA cases					Controls						
	No. of subjects	Allele G	Allele T	Allele A	MAF	No. of subjects	Allele G	Allele T	Allele A	MAF	OR (95% CI)	*P*
STAT4 rs7574865												
Caucasian populations												
This study	854	1,276	432		0.25	866	1,348	384		0.22	1.19 (1.01–1.40)	0.031
Remmers et al 2007												
NARAC	606	872	340		0.28	1,309	2,042	576		0.22	1.38 (1.18–1.62)	4.4×10^{-5}
Replication study, US	1,013	1,499	527		0.26	1,326	2,069	583		0.22	1.25 (1.09–1.43)	1.3×10^{-3}
EIRA	1,529	2,293	765		0.25	881	1,374	388		0.22	1.18 (1.03–1.36)	0.018
Barton et al 2008												
WTCCC	1,858	2,835	881		0.24	2,934	4,580	1,288		0.22	1.11 (1.00–1.22)	0.045
Replication study, UK	3,399	5,140	1,658		0.24	3,024	4,744	1,304		0.22	1.17 (1.08–1.28)	1.4×10^{-4}
Orozco et al 2008												
Spanish	923	1,389	457		0.24	1,296	2,054	538		0.21	1.26 (1.09–1.45)	1.6×10^{-3}
Dutch	876	1,319	433		0.25	893	1,399	387		0.22	1.19 (1.01–1.39)	0.031
Swedish	273	388	158		0.29	285	438	132		0.23	1.35 (1.03–1.77)	0.028
Palomino-Morales et al 2008	257	316	198		0.38	410	562	258		0.31	1.36 (1.08–1.73)	8.1×10^{-3}
Zervou et al 2008	311	451	171		0.27	344	574	114		0.17	1.91 (1.46–2.49)	1.7×10^{-6}
Pooled Caucasian	11,899					13,568					1.24 (1.17–1.33)	1.66×10^{-11}
East Asian populations												
Lee et al 2007	1,032	1,269	795		0.38	908	1,215	601		0.33	1.27 (1.11–1.45)	4.5×10^{-4}
Kobayashi et al 2008												
Tokyo, Japan	1,481	1,870	1,092		0.37	745	1,026	464		0.31	1.29 (1.13–1.48)	1.6×10^{-4}
Biobank Project, Japan	1,109	1,396	822		0.37	938	1,295	581		0.31	1.31 (1.15–1.50)	4.3×10^{-5}
Tokushima, Japan	941	1,178	704		0.37	500	662	338		0.34	1.17 (0.99–1.38)	0.055
Pooled East Asian	4,563					3,091					1.27 (1.18–1.36)	1.4×10^{-11}
Pooled Caucasian + East Asian	16,462					16,659					1.24 (1.19–1.31)	$<1 \times 10^{-15}$
IL2/IL21 rs6822844												
Caucasian populations												
This study	877	1,469	285		0.16	866	1,407	325		0.19	0.84 (0.70–1.00)	0.051
Zhernakova et al 2007	1,012	1,739	285		0.14	924	1,506	342		0.19	0.72 (0.61–0.86)	1.9×10^{-4}
Pooled Caucasian	1,889					1,790					0.78 (0.69–0.88)	5.6×10^{-5}
CTLA4 rs3087243												
Caucasian populations												
This study	867	1,005		729	0.42	863	941		785	0.45	0.87 (0.76–1.00)	0.045
Plenge et al 2005												
EIRA	1,505	1,870		1,140	0.38	878	1,070		686	0.39	0.95 (0.84–1.08)	0.410
NARAC	828	1,003		653	0.39	845	934		756	0.45	0.80 (0.70–0.93)	0.001
Barton et al 2004	719	820		618	0.43	755	848		662	0.44	0.97 (0.83–1.12)	0.636
Orozco et al 2004	433	432		434	0.50	398	401		395	0.50	1.02 (0.84–1.24)	0.841
Zhernakova et al 2005	153	173		133	0.43	900	959		841	0.47	0.88 (0.68–1.13)	0.291
Pooled Caucasian	4,505					4,639					0.91 (0.85–0.96)	0.0018
East Asian populations												
Lei et al 2005	326	449		203	0.31	250	305		195	0.39	0.71 (0.55–0.91)	0.005
Tsukahara et al 2008	1,498	2,284		712	0.24	441	653		229	0.26	0.89 (0.75–1.06)	0.181
Pooled East Asian	1,824					691					0.82 (0.72–0.95)	0.007
Pooled Caucasian + East Asian	6,329					5,330					0.89 (0.85–0.95)	8.3×10^{-5}

* Values for alleles G, T, and A are the allele frequency. Rheumatoid arthritis (RA) cases and controls were compared using a fixed-effects (Mantel-Haenszel) meta-analysis. No significant heterogeneity was observed among the studies. MAF = minor allele frequency; OR = odds ratio; 95% CI = 95% confidence interval; NARAC = North American Rheumatoid Arthritis Consortium; EIRA = Epidemiological Investigation of Rheumatoid Arthritis; WTCCC = Wellcome Trust Case Control Consortium.

Table 3. Meta-analysis of *STAT4* (rs7574865), *IL2/IL21* (rs6822844), and *CTLA4* (rs3087243) stratified by autoantibody status*

Study	No. of subjects	Allele G	T	A	MAF	Cases vs. controls OR (95% CI)	*P*	ACPA+ vs. ACPA− OR (95% CI)	*P*	RF+ vs. RF− OR (95% CI)	*P*
STAT4 rs7574865											
Caucasian populations											
This study								1.03 (0.77–1.38)	0.819		
ACPA+	318	478	158		0.25	1.16 (0.93–1.44)	0.170				
ACPA−	218	325	111		0.25	1.20 (0.93–1.54)	0.144				
Barton et al 2008								1.01 (0.86–1.91)	0.880		
ACPA+	1,211	1,823	599		0.25	1.20 (1.07–1.34)	0.001				
ACPA−	617	926	308		0.25	1.21 (1.05–1.40)	0.009				
Orozco et al 2008								0.94 (0.69–1.27)	0.672		
ACPA+	288	421	155		0.27	1.41 (1.14–1.74)	0.001				
ACPA−	187	278	96		0.27	1.32 (1.02–1.71)	0.030				
Pooled Caucasian								1.00 (0.89–1.14)	0.97		
ACPA+	1,817					1.22 (1.12–1.33)	8.1×10^{-6}				
ACPA−	1,022					1.23 (1.10–1.37)	2.8×10^{-4}				
East Asian populations											
Lee et al 2007								1.00 (0.74–1.35)	0.985		
ACPA+	612	749	475		0.39	1.01 (0.87–1.17)	0.869				
ACPA−	111	136	86		0.39	1.01 (0.75–1.35)	0.949				
Pooled Caucasian + East Asian								1.00 (0.89–1.12)	0.98		
IL2/IL21 rs6822844											
Caucasian populations											
This study										1.09 (0.80–1.48)	0.577
RF+	487	811	163		0.17	0.87 (0.70–1.07)	0.188				
RF−	250	422	78		0.16	0.80 (0.81–1.06)	0.105				
Zhernakova et al 2007										1.22 (0.82–1.83)	0.306
RF+	664	1,143	185		0.14	0.71 (0.58–0.87)	0.0006				
RF−	112	187	37		0.17	0.87 (0.59–1.28)	0.467				
Pooled Caucasian										1.02 (0.81–1.29)	0.86
RF+	1,151					0.78 (0.68–0.90)	6.9×10^{-4}				
RF−	362					0.82 (0.66–1.03)	0.083				
CTLA4 rs3087243											
Caucasian populations											
This study								0.88 (0.69–1.13)	0.324		
ACPA+	322	387		257	0.40	0.80 (0.66–0.96)	0.015				
ACPA−	225	257		193	0.43	0.90 (0.73–1.11)	0.325				
Plenge et al 2005		644									
EIRA		450						0.84 (0.73–0.98)	0.03		
ACPA+	904	1,152		656	0.36	0.89 (0.78–1.02)	0.08				
ACPA−	581	694		468	0.40	1.05 (0.90–1.22)	0.51				
NARAC								0.87 (0.67–1.11)	0.26		
ACPA+	572	697		447	0.39	0.79 (0.68–0.92)	0.003				
ACPA−	161	185		137	0.43	0.91 (0.72–1.16)	0.48				
Karlson et al 2008								0.88 (0.70–1.11)	0.293		
ACPA+	436	514		358	0.41						
ACPA−	220	246		194	0.44						
Pooled Caucasian								0.86 (0.78–0.96)	0.0047		
ACPA+	2,234					0.83 (0.76–0.91)	5.4×10^{-5}				
ACPA−	1,187					0.98 (0.88–1.09)	0.709				

* Values for alleles G, T, and A are the allele frequency. Data were compared using a fixed-effects (Mantel-Haenszel) meta-analysis. No significant heterogeneity was observed. A significant ($P < 0.05$) association could be observed only for *CTLA4* rs3087243 in anti–citrullinated protein antibody (ACPA)–positive patients. RF = rheumatoid factor (see Table 2 for other definitions).

Discussion

In the present study, 2 genetic risk factors for RA were replicated in an independent Dutch population, with a third genetic risk factor showing a clear trend toward association. All 3 loci were further confirmed in a well-powered meta-analysis. Interestingly, these polymorphisms have been described in several autoimmune diseases, varying from type 1 diabetes to systemic lupus erythematosus (3,4), which further emphasizes their role in autoimmunity. In RA, previous studies have suggested that genetic risk factors predispose to specific subsets of the disease, characterized by autoantibody status. For example, both the *HLA* shared epitope and *PTPN22* loci have been shown to be associated with a clear predisposition to ACPA-positive disease only. At a biologic level, classifying these genetic risk factors will ultimately enable a better understanding of the disease processes involved. Although *STAT4*, *IL2*/*IL21*, and *CTLA4* have been found in association with ACPA-positive disease (20), they have not been investigated extensively in autoantibody-negative patients.

In this study, we did not observe a difference in effect size between the autoantibody strata for either *STAT4* or *IL2*/*IL21*. Our results are consistent with recent findings for *STAT4.* Zhernakova and colleagues (4) have shown an association of *IL2*/*IL21* in RF-positive patients as compared with controls, but no conclusive difference could be established between the 2 subgroups. In support of these findings, our combined data sets also indicated that *IL2*/*IL21* predisposes individuals to both autoantibody-positive and autoantibody-negative disease. However, additional replication in independent cohorts will still be necessary to tease apart the precise role of *IL2*/*IL21* in these disease subsets.

The results from several studies previously suggested that *CTLA4* is associated with RA in an autoantibody-dependent manner. However, we provided, for the first time, conclusive evidence that *CTLA4* is associated with ACPA-positive RA, but not with ACPA-negative RA. In conclusion, this study provides independent replication of an association of *STAT4, IL2*/*IL21,* and *CTLA4* with RA, as well as substantial evidence of the involvement of *CTLA4* in ACPA-positive disease only, as compared with the involvement of *STAT4* and *IL2*/ *IL21*, which predisposes to both disease subsets.

References

(1) Newton JL, Harney SM, Wordsworth BP, Brown MA. A review of the MHC genetics of rheumatoid arthritis. Genes Immun 2004;5: 151–7.

(2) Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–78.

(3) Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007;357:977–86.

(4) Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, Franke B, et al. Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. Am J Hum Genet 2007;81:1284–8.

(5) Van der Helm-van Mil AH, Huizinga TW, de Vries RR, Toes RE. Emerging patterns of risk factor make-up enable subclassification of rheumatoid arthritis. Arthritis Rheum 2007;56:1728–35.

(6) Barton A, Thomson W, Ke X, Eyre S, Hinks A, Bowes J, et al. Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility. Hum Mol Genet 2008;17: 2274–9.

(7) Kobayashi S, Ikari K, Kaneko H, Kochi Y, Yamamoto K, Shimane K, et al. Association of STAT4 with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in the Japanese population. Arthritis Rheum 2008;58:1940–6.

(8) Lee HS, Remmers EF, Le JM, Kastner DL, Bae SC, Gregersen PK. Association of STAT4 with rheumatoid arthritis in the Korean population. Mol Med 2007;13:455–60.

(9) Orozco G, Alizadeh BZ, Gado-Vega AM, Gonzalez-Gay MA, Balsa A, Pascual-Salcedo D, et al. Association of STAT4 with rheumatoid arthritis: a replication study in three European populations. Arthritis Rheum 2008;58:1974–80.

(10) Palomino-Morales RJ, Rojas-Villarraga A, Gonzalez CI, Ramirez G, Anaya JM, Martin J. STAT4 but not TRAF1/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians. Genes Immun 2008;9: 379–82.

(11) Zervou MI, Sidiropoulos P, Petraki E, Vazgiourakis V, Krasoudaki E, Raptopoulou A, et al. Association of a TRAF1 and a STAT4 gene polymorphism with increased risk for rheumatoid arthritis in a genetically homogeneous population. Hum Immunol 2008;69:567–71.

(12) Barton A, Jury F, Eyre S, Bowes J, Hinks A, Ward D, et al. Haplotype analysis in simplex families and novel analytic approaches in a case-control cohort reveal no evidence of association of the CTLA-4 gene with rheumatoid arthritis. Arthritis Rheum 2004;50:748–52.

(13) Lei C, Dongqing Z, Yeqing S, Oaks MK, Lishan C, Jianzhong J, et al. Association of the CTLA-4 gene with rheumatoid arthritis in Chinese Han population. Eur J Hum Genet 2005;13:823–8. 14. Orozco G, Torres B, Nunez-Roldan A, Gonzalez-Escribano MF, Martin J. Cytotoxic T-lymphocyte antigen-4-CT60 polymorphism in rheumatoid arthritis. Tissue Antigens 2004;64:667–70.

(14) Plenge RM, Padyukov L, Remmers EF, Purcell S, Lee AT, Karlson EW, et al. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. Am J Hum Genet 2005;77:1044–60.

(15) Tsukahara S, Iwamoto T, Ikari K, Inoue E, Tomatsu T, Hara M, et al. CTLA-4 CT60 polymorphism is not an independent genetic risk marker of rheumatoid arthritis in a Japanese population. Ann Rheum Dis 2008;67:428–9.

(16) Zhernakova A, Eerligh P, Barrera P, Wesoly JZ, Huizinga TW, Roep BO, et al. CTLA4 is differentially associated with autoimmune diseases in the Dutch population. Hum Genet 2005;118: 58–66.

(17) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

(18) Kurreeman FA, Padyukov L, Marques RB, Schrodi SJ, Seddighzadeh M, Stoeken-Rijsbergen G, et al. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS Med 2007;4:e278.

(19) Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP, et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 2008;40:1216–23.

Chapter 13

Association of the Autoimmunity Locus 4q27 With Juvenile Idiopathic Arthritis

H.M. Albers
F.A.S. Kurreeman
G. Stoeken-Rijsbergen
D.M.C. Brinkman
S.S.M. Kamphuis
M.A.J. van Rossum
H.J. Girschick
C. Wouters
R.K. Saurenmann
E. Hoppenreijs
P. Slagboom
J.J. Houwing-Duistermaat
W. Verduijn
T.W.J. Huizinga
R. ten Cate
R.E.M. Toes
M.W. Schilham

Arthritis Rheum. 60, 901-4 (2009).

Abstract

Objective

Juvenile idiopathic arthritis (JIA) is characterized by chronic arthritis and an autoimmune etiology. In several autoimmune diseases, including rheumatoid arthritis (RA), an association with the 4q27 locus has been reported. We undertook this study to investigate the possible role of the 4q27 locus in JIA.

Methods

A case-control association study was conducted, with a total of 655 Caucasian JIA patients and 791 healthy controls divided into 2 independent sample sets. The rs6822844 marker in the 4q27 locus was genotyped.

Results

In the first and larger sample set, a 5% decrease in T allele frequency was observed in patients compared with controls (allelic odds ratio [OR] 0.72 [95% confidence interval 0.55–0.95], *P* =0.019), and in the second set, a 3% decrease was observed (allelic OR 0.81 [95% confidence interval 0.61–1.09], *P* =0.169). The combined data set generated an OR of 0.76 (95% confidence interval 0.62–0.93, P=7.08 x10^{-3}). When the different JIA subtypes were analyzed individually, significant decreases were seen in the subtypes with a polyarticular course of disease (extended oligoarthritis [*P* = 0.019] and rheumatoid factor–negative polyarthritis [*P* = 0.038]).

Conclusion

Our findings suggest that the 4q27 locus, previously reported to be associated with RA, type 1 diabetes mellitus, celiac disease, and psoriatic arthritis, is also associated with susceptibility to JIA.

Introduction

Juvenile idiopathic arthritis (JIA) is a group of heterogeneous disorders characterized by chronic arthritis diagnosed in children younger than 16 years of age (1). Seven different subtypes of JIA can be distinguished according to the International League of Associations for Rheumatology (ILAR) classification (2). The subtypes oligoarthritis (both persistent and extended) and rheumatoid factor (RF)-negative polyarthritis are considered the most homogeneous subtypes, with shared phenotypic features. Systemic JIA has a more distinct phenotype resembling an autoinflammatory syndrome. Although the precise etiology is still unknown, JIA is considered to be an autoimmune disease.

Genetic studies in autoimmune diseases have revealed the presence of shared common autoimmune susceptibility loci (3). In JIA, associations with the major histocompatibility complex locus, PTPN22, and TRAF1/C5 have been described (4–6). The 4q27 locus, a region of strong linkage disequilibrium that includes IL2 and IL21 (the genes encoding interleukin-2 [IL-2] and IL-21, respectively), has been associated with celiac disease, rheumatoid arthritis (RA), type 1 diabetes mellitus (type 1 DM), and psoriatic arthritis (7–9). To determine whether the 4q27 locus is also associated with JIA, we genotyped rs6822844 (which can be used as a proxy for the 4q27 haplotype that is associated with autoimmune disease [8]) in JIA patients and controls.

Patients & Methods

Patient population

A case-control association study was conducted in 2 independent sample sets. These sample sets consisted of Caucasian JIA patients (recruited from pediatric rheumatology centers in The Netherlands [n =327], Belgium [n= 96], Germany [n=95], and Switzerland [n =137]) and healthy Dutch adult controls, who were part of the DNA panel of the Immunogenetics and Transplantation Immunology section at Leiden University Medical Center and consisted of randomly selected unrelated individuals and blood donors.

The first sample set consisted of 328 JIA patients and 465 healthy controls, and the second sample set consisted of 327 JIA patients and 326 controls. In all patients (69% female and 31% male), JIA was diagnosed according to the ILAR revised classification criteria (2). The patients with JIA who were selected for inclusion in the study were primarily those with the oligoarthritis (both persistent and extended) or RF-negative polyarthritis subtype, because of their homogeneous phenotypes.

Forty-four percent of the patients in the overall JIA patient group had persistent oligoarticular disease, 13% had extended oligoarticular disease, 24% had RF-negative polyarticular disease, 3% had RF-positive polyarticular disease, 11% had systemic JIA, and 5% had other JIA subtypes. Because of the small sample size of patients with RF-positive polyarticular JIA and patients with types of JIA categorized as "other," these groups were excluded from the subtype analysis described below. All patients were of self- or parent-reported European Caucasian ethnicity. Written informed consent was obtained from all patients and/or their parents, and the institutional review boards from all participating centers approved the study.

DNA and Genotyping

To test for an association of the 4q27 region with JIA, rs6822844 was typed in both sample sets. DNA was collected through either a blood sample (12% of JIA patients and all controls) or a mouth swab (88% of JIA patients). Genotyping was performed using high-throughput MassArray matrix-assisted laser desorption ionization time of flight mass spectrometry, according to the protocols recommended by the manufacturer (Sequenom, San Diego, CA).

Each of the 384-well plates contained 8 positive controls (DNA obtained from the Centre d'Etude du Polymorphisme Humain, Paris, France), 8 negative controls, and 10% duplicates. The error rate was <1%. Random genotyping failure occurred in 3% of JIA patients and 1% of controls, decreasing the total number of subjects in the analyses to 635 patients and 783 controls.

Statistical analysis

Because of the adherence to an additive model and the lack of evidence for a recessive model, we compared cases and controls using an allelic odds ratio (OR) and 95% confidence interval (95% CI). The common OR of the 2 independent sample sets combined was calculated using the Mantel-Haenszel test. There was no heterogeneity between the 2 sample sets (P >0.05 by Breslow-Day test). Case and control genotype frequencies did not deviate from Hardy-Weinberg equilibrium. All statistical analyses were performed with SPSS, version 14.0 (SPSS, Chicago, IL). P values less than 0.05 were considered significant.

Results

In the first and larger sample set, the T allele frequency was significantly decreased, from 20% in controls to 15% in JIA patients (allelic OR 0.72 [95% CI 0.55–0.95], $P = 0.019$) (Table 1). A similar decrease in T allele frequency (from 18% in controls to 15% in patients) was observed in the second sample set, although it did not reach statistical significance (allelic OR 0.81 [95% CI 0.61–1.09], $P = 0.169$). The common OR in sample sets 1 and 2 combined demonstrated a positive association with JIA ($P = 7.08 \times 10^{-3}$).

Table 1. Genotype and allele frequencies (rs6822844) in JIA patients and controls in 2 independent sample sets*

	Genotype frequency, no. (%)			T allele frequency, no. (%)	Allelic OR (95% CI)	*P*
	GG	GT	TT			
Set 1						
Controls (n = 460)	293 (63.7)	152 (33.0)	15 (3.3)	182 (19.8)	–	–
Patients (n = 311)	224 (72.0)	80 (25.7)	7 (2.3)	94 (15.1)	0.72 (0.55–0.95)	0.019
Set 2						
Controls (n = 323)	218 (67.5)	95 (29.4)	10 (3.1)	115 (17.8)	–	–
Patients (n = 324)	233 (71.9)	85 (26.2)	6 (1.9)	97 (15.0)	0.81 (0.61–1.09)	0.169
Combined data sets						
Controls (n = 783)	511 (65.3)	247 (31.5)	25 (3.2)	297 (19.0)	–	–
Patients (n = 635)	457 (72.0)	165 (26.0)	13 (2.0)	191 (15.0)	0.76 (0.62–0.93)†	7.08×10^{-3}

* JIA = juvenile idiopathic arthritis; 95% CI = 95% confidence interval.
† Mantel-Haenszel odds ratio (OR).

Because of the importance of investigating genetic risk factors in homogeneous, well-defined phenotypic groups, we also analyzed association in the selected JIA subtypes (Table 2). Although a trend toward a decreased T allele frequency was observed among patients with persistent oligoarthritis and systemic JIA, only the subtypes with a polyarticular course of disease (extended oligoarthritis and RF-negative polyarthritis) showed a significant decrease in T allele frequency ($P = 0.019$ and $P = 0.038$, respectively).

Table 2. Analysis of allele frequencies in patients with selected subtypes of JIA and in controls*

	T allele	Allelic OR (95% CI)	*P*
JIA subtype†			
Persistent oligoarthritis (n = 275)	0.165	0.85 (0.65–1.10)	0.207
Extended oligoarthritis (n = 83)	0.114	0.55 (0.34–0.91)	0.019
RF-negative polyarthritis (n = 151)	0.139	0.69 (0.49–0.98)	0.038
Systemic JIA (n = 69)	0.152	0.77 (0.47–1.24)	0.280
Control (n = 783)	0.190	–	–

* RF = rheumatoid factor (see Table 1 for other definitions).
† Diagnosed according to the International League of Associations for Rheumatology revised classification criteria (2).

Discussion

This study demonstrates, for the first time, a positive association of the 4q27 locus (rs6822844) with JIA; a JIA-protective effect of the T allele was observed. When the different JIA subtypes were tested individually, a decrease in T allele frequency was observed in all subtypes. Interestingly, this decrease was significant only in the JIA subtypes that share a polyarticular phenotype. Similar observations have been made regarding the recently identified association with the *TRAF1/C5* region (4), indicating a common genetic constitution underlying the polyarticular JIA phenotype.

In studies of patients originating from different European countries, population stratification cannot be ruled out completely. However, no significant variance in allele frequency in control populations from Western Europe has been reported (7,8), and the allele frequency observed in our control population was very similar to previously reported frequencies. In addition, when allele frequencies in patients originating from the various European countries were compared, no significant difference was found (P = 0.77). Moreover, the degree of underrepresentation of the minor T allele and the 4q27 effect size were similar to those described in studies of other autoimmune diseases, such as RA and type 1 DM (8).

The 4q27 locus consists of a large region of strong linkage disequilibrium encoding the genes *KIAA1109*, *TENR*, *IL2,* and *IL21*. Both *IL2* and *IL21* are likely candidates for association with susceptibility to JIA, since both cytokines are involved in immune activation and regulation pathways.

The IL-2 pathway, in which the interaction between IL2 and IL2 receptor α is central, is involved in T cell proliferation and regulation (10). In mice deficient in *IL2,* T cells have impaired proliferation and effector function in vitro, and lethal autoimmunity develops (11). In addition to the 4q27 locus, the *IL2RA* locus has been associated with several autoimmune diseases (12,13), indicating an important role of the IL2 pathway in immune regulation and maintenance of self-tolerance.

IL21 is also involved in a wide range of immunologic processes. It appears to play a role in autoimmunity by influencing the cellular immune response through both inhibition of suppression by CD4+ regulatory T cells and generation of Th17 cells, as well as by influencing the humoral response (14).

Further analysis of the immunologic pathways involved in JIA may be helpful in identifying the causal gene in this locus. Moreover, sequencing, fine-mapping, and extensive testing of variants of this region will be required to narrow down the region of association and identify the associated gene. Functional testing of all the linked variants associated with disease is also needed.

In conclusion, like RA, type 1 DM, celiac disease, and psoriatic arthritis, JIA is associated with the 4q27 locus. The identification of the 4q27 locus as a risk factor for JIA contributes to the accumulated evidence that one of the genes in this region plays a role in autoimmune diseases in general. In addition, our data indicate that the 4q27 locus contributes to genetic susceptibility to JIA, warranting further research into the biologic pathways that would help to explain this association.

References

(1) Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007;369:767–78.

(2) Petty RE, Southwood TR, Manners P, Baum J, Glass DN,Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.

(3) Phelan JD, Thompson SD, Glass DN. Susceptibility to JRA/JIA: complementing general autoimmune and arthritis traits. Genes Immun 2006;7:1–10.

(4) Albers HM, Kurreeman FA, Houwing-Duistermaat JJ, Brinkman DM, Kamphuis SS, Girschick HJ, et al. The TRAF1/C5 region is a risk factor for polyarthritis in juvenile idiopathic arthritis. Ann Rheum Dis 2008;67:1578–80.

(5) Hinks A, Worthington J, Thomson W. The association of PTPN22 with rheumatoid arthritis and juvenile idiopathic arthritis. Rheumatology (Oxford) 2006;45:365–8.

(6) Thompson SD, Moroldo MB, Guyer L, Ryan M, Tombragel EM, Shear ES, et al. A genome-wide scan for juvenile rheumatoid arthritis in affected sibpair families provides evidence of linkage. Arthritis Rheum 2004;50:2920–30.

(7) Van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, Inouye M, et al. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat Genet 2007;39:827–9.

(8) Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, Franke B, et al. Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. Am J Hum Genet 2007;81:1284-8.

(9) Liu Y, Helms C, Liao W, Zaba LC, Duan S, Gardner J, et al. A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. PLoS Genet 2008;4:e1000041.

(10) Malek TR. The biology of interleukin-2. Annu Rev Immunol 2008;26:453–79.

(11) Sadlack B, Lohler J, Schorle H, Klebb G, Haber H, Sickel E, et al. Generalized autoimmune disease in interleukin-2-deficient mice is triggered by an uncontrolled activation and proliferation of CD4_ T cells. Eur J Immunol 1995;25:3053–9.

(12) Brand OJ, Lowe CE, Heward JM, Franklyn JA, Cooper JD, Todd JA, et al. Association of the interleukin-2 receptor _ (IL-2R_)/CD25 gene region with Graves' disease using a multilocus test and tag SNPs. Clin Endocrinol (Oxf) 2007;66:508–12.

(13) Vella A, Cooper JD, Lowe CE, Walker N, Nutland S, Widmer B, et al. Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. Am J Hum Genet 2005;76:773–9.

(14) Ettinger R, Kuchen S, Lipsky PE. The role of IL-21 in regulating B-cell function in health and disease. Immunol Rev 2008;223: 60–86.

Chapter 14

Association of a Haplotype in the Promoter Region of the Interferon regulatory Factor 5 Gene With Rheumatoid Arthritis

S. Sigurdsson
L. Padyukov
F.A.S. Kurreeman
U. Liljedahl
A.C. Wiman
L. Alfredsson
R.E.M. Toes
J. Rönnelid
L. Klareskog
T.W.J. Huizinga
G. Alm
A.C. Syvänen
L. Rönnblom

Arthritis Rheum. 56, 2202-2210 (2007)

Abstract

Objective

To determine whether genetic variants of the interferon regulatory factor 5 (IRF-5) and Tyk-2 genes are associated with rheumatoid arthritis (RA).

Methods

Five single-nucleotide polymorphisms (SNPs) in *IRF5* and 3 SNPs in *Tyk2* were analyzed in a Swedish cohort of 1,530 patients with RA and 881 controls. A replication study was performed in a Dutch cohort of 387 patients with RA and 181 controls. All patient sera were tested for the presence of autoantibodies against cyclic citrullinated peptides (anti-CCP).

Results

Four of the 5 SNPs located in the 5' region of *IRF5* were associated with RA, while no association was observed with the *Tyk2* SNPs. The minor alleles of 3 of the *IRF5* SNPs, which were in linkage disequilibrium and formed a relatively common haplotype with a frequency of ~0.33, appeared to confer protection against RA. Although these disease associations were seen in the entire patient group, they were mainly found in RA patients who were negative for anti-CCP. A suggestive association of *IRF5* SNPs with anti-CCP–negative RA was also observed in the Dutch cohort.

Conclusion

Given the fact that anti-CCP–negative RA differs from anti-CCP–positive RA with respect to genetic and environmental risk factor profiles, our results indicate that genetic variants of *IRF5* contribute to a unique disease etiology and pathogenesis in anti-CCP–negative RA.

Introduction

The type I interferons (IFNs) comprise a large family of cytokines that are typically produced during viral infections, mainly by plasmacytoid dendritic cells (1). In addition to direct antiviral effects, type I IFNs have many important immunomodulatory functions (1). For instance, they cause maturation of dendritic cells, promote T cell activation, and stimulate B cell development and production of antibodies. Data from previous studies indicate that the type I IFN system plays a pivotal role when self tolerance is broken and autoimmune reactions develop (2). Accordingly, the type I IFN system is activated in many autoimmune diseases, including systemic lupus erythematosus (SLE) (2), pri-mary Sjogren's syndrome (3), psoriasis (4), polymyositis (5), and type 1 diabetes mellitus (6).

In a previous report we described a strong association between SLE and 2 key genes from the type I IFN signaling system, namely the Tyk-2 and interferon regulatory factor (IRF-5) genes (7). The association of the single-nucleotide polymorphisms (SNPs) in the IRF-5 gene with SLE has recently been convincingly replicated in 4 independent case–control studies (8). A causative role of type I IFN in the initiation and maintenance of autoimmunity is suggested by the finding that up to 19% of IFN=-treated patients with a malignant disease ultimately developed an autoimmune disorder (2). Development of rheumatoid arthritis (RA [MIM no. 180300]) during IFNα treatment has been reported (9), providing evidence that the type I IFN system may be involved in the pathogenesis of this disease.

RA is a common autoimmune disease characterized by chronic arthritis with progressive joint destruction, and if left untreated RA can lead to severe disability. RA affects women 2–3 times more frequently than it affects men, and both genetic and environmental components are involved in the etiopathogenesis of the disease (10). A majority of RA patients are positive for autoantibodies against the Fc portion of the IgG molecule (rheumatoid factor [RF]). The presence of RF in patients correlates with a severe form of RA (11), but RF can occasionally be detected in normal healthy individuals and in patients with conditions other than RA (12).

It has been shown that RA patients also produce antibodies against cyclic citrullinated peptides (anti-CCP), and that anti-CCP antibodies are rarely detectable in healthy individuals (13). They are thus a more specific disease marker for RA than is RF. Anti-CCP antibodies can be detected years before the appearance of any clinical RA symptoms (14), and they are prognostic indicators of a greater degree of inflammation and more destructive disease (15). Recent data on gene–environment interactions also show that the major environmental risk factor, smoking, is associated with anti-CCP–positive RA, but not with anti-CCP–negative RA (10,16). Furthermore, anti-CCP antibodies display a strong association with specific alleles at the HLA–DRB1 gene (10), which are collectively known as shared epitope alleles. In contrast, the HLA–DR3 allele is associated with anti-CCP–negative RA, which suggests that the syndrome called RA may consist of at least 2 distinct diseases (anti-CCP positive and anti-CCP negative), each with different risk factors and pathogenetic pathways (17). Genetic variants of several non–major histocompatibility complex (MHC) genes have also shown an association with RA (18), and among them the most convincing evidence for association has been observed for variants of the *PTPN22* (MIM no. 600716) (19), *PADI4* (MIM no. 605347) (20), and *CTLA4* (MIM no. 123890) (21,22) genes.

The *PTPN22* allele, which has shown the most consistent association with RA in previously published studies, is associated with anti-CCP–positive RA (18). So far, no genetic risk factor outside the HLA locus has been identified for anti-CCP–negative disease.

The role of the type I IFN system in RA has not been extensively investigated, but expression of IFNα (23) and the type I IFN–inducible protein myxovirus resistance protein A (24) has been detected in RA synovial tissue, along with increased levels of IFNα in synovial fluid (25). Furthermore, plasmacytoid dendritic cells are recruited to inflamed synovium (26), perhaps due to increased levels of production of chemoattractants for these cells in the synovial fluid (24). Moreover, the cytokines interleukin-6 (IL-6), IL-10, IL-12, and tumor necrosis factor α (TNFα), which are connected to the type I IFN pathway, have all been suggested to be involved in the disease process in RA (27).

The JAK Tyk-2 is crucial for signaling via the type I IFN receptor. Tyk-2 also interacts with the receptors for several other cytokines, such as IL-6, IL-10, and IL-12. A possible role of Tyk-2 in RA is more directly indicated by the finding that Tyk-2–deficient mice are resistant to experimental arthritis (28). IRF-5 is a transcription factor that is activated by Toll-like receptor 7 (TLR-7), TLR-8, and TLR-9, which are involved in induction of IFN= genes in human cells (29,30). IRF-5 is constitutively expressed in cells of the immune system, including plasmacytoid dendritic cells (31), and its expression is regulated by type I IFN (32).

Given the potential role of the type I IFN system in RA, we hypothesized that genetic variants of *Tyk2* and *IRF5* could be associated with RA, by analogy with our previous findings in SLE (7). To test this hypothesis, we genotyped 5 SNPs in the IRF-5 gene and 3 SNPs in the Tyk-2 gene in a large cohort of RA patients and controls from Sweden. We found that 4 of the 5 *IRF5* SNPs and a commonly observed haplotype formed by the rare alleles of these SNPs were strongly associated with RA, and that the evidence of association was mainly found in anti-CCP–negative RA. We replicated the association between *IRF5* and anti-CCP–negative RA in an unrelated cohort of Dutch patients.

Patients & Methods

Samples

The Swedish RA patients and controls were from the Epidemiological Investigation of Rheumatoid Arthritis (EIRA), a population-based case–control study of incident cases of RA. The EIRA included subjects ages 18–70 years who were living in the central and southern parts of Sweden during 1996–2003. Briefly, all potential cases were examined and diagnosed by a rheumatologist at a study unit. All rheumatology units linked to the public health care system, as well as almost all of the few existing privately run rheumatology units, in the relevant geographic area participated in the study. In total there were 19 reporting clinics, 15 of which were Early Arthritis Clinics. Initially, some centers reported cases that did not satisfy the criteria for RA, in order to enable investigations of undifferentiated arthritis, but these subjects were eventually excluded from the study.

For each potential case, a control subject was randomly selected from the study base, with consideration of age, sex, and area of residence. Controls were selected using the Swedish national population register, which is continuously updated. If a proposed control declined to participate, was not traceable, or was reported as having RA, a new control was selected using the same criteria. Controls selected to match cases who were eventually excluded because they did not fulfill study criteria remained in the study.

In the present study we used part of the EIRA cohort consisting of 1,530 newly diagnosed patients who fulfilled the 1987 American College of Rheumatology (formerly, the American Rheumatism Association) criteria for RA (33), and 881 controls who answered a questionnaire and provided a blood sample. Among the cases, the mean duration between estimated disease onset and study entry was 10 months. Of all participants in the EIRA study, 97% were Caucasian. Seventeen controls who were found to be positive for anti-CCP antibodies and 3 controls with unknown anti-CCP status were excluded from the analysis.

The replication cohort consisted of RA patients and controls from the Leiden Early Arthritis Clinic, as previously described in detail (34). Briefly, in 1993, an Early Arthritis Clinic was established at the Department of Rheumatology of the Leiden University Medical Center. General practitioners in an area with ~300,000 inhabitants referred patients directly to the clinic when arthritis was suspected. All patients were evaluated at the Early Arthritis Clinic within 2 weeks of referral. Patients were included in the study only if the duration of symptoms was <2 years, and the mean duration between estimated disease onset and study entry was 7 months. Diagnoses were made according to international classification criteria, as previously described (35).

Patients from the Early Arthritis Clinic who were diagnosed as having definite RA, and from whom DNA was available, were included in the present study (n = 387). Other anti-CCP–negative and RF-negative patients from the Early Arthritis Clinic with diagnoses such as gout, pseudogout, viral or bacterial reactive arthritis, posttraumatic osteoarthritis, Lyme arthritis, or paraneoplastic arthritis made up the control group (n = 181). Patients with undifferentiated arthritis, psoriatic arthritis, or SLE were excluded from the control group. All subjects in the case and control groups were Caucasian.

All subjects in both cohorts gave informed consent, and the local ethics committee approved the study. DNA was isolated from EDTA-treated blood by a standard desalting method.

Autoantibody Analysis

Patient sera obtained from the EIRA cohort at the time of diagnosis were examined for RF by nephelometry and for anti-CCP antibodies by enzyme-linked immunosorbent assay (ELISA) (second-generation test; Euro-Diagnostica, Arnhem, The Netherlands). Antinuclear antibodies (ANAs) were measured in 194 anti-CCP–positive and 191 anti-CCP–negative patients with RA using indirect immunofluorescence on HEp-2 cells (Bio-Rad, Stockholm, Sweden) in a 1:200 screening dilution with a secondary antibody directed against the γ-chain of IgG (Dako, Glostrup, Denmark). All sera were interpreted blindly in parallel, by the same investigator (JR). Using this procedure, ANAs were detected in sera from 5 (5%) of 100 healthy controls. In the Early Arthritis Clinic study, serum anti-CCP antibodies were assessed with a commercial ELISA (Immunoscan RA [Mark 2]; Euro-Diagnostica). Anti-CCP antibodies were measured in serum collected within 4 months after the first visit (in 94% of the patients) or, when serum was not available within this period, in the first stored serum sample available thereafter.

Genotyping

In the Swedish EIRA cohort we genotyped 4 SNPs in the promoter and first intron of the IRF-5 gene (no. rs729302, no. rs2004640, no. rs752637, and no. rs3807306) and 2 SNPs in protein-coding exons of the Tyk-2 gene (no. rs12720356 and no. rs2304256) for which we had observed strong or suggestive signals for joint linkage and association with SLE in a previous study (7). We also included in the analysis an additional SNP in the promoter region of *IRF5* (no. rs3757385) and a *Tyk2* SNP (no. rs91755) located in intron 9 between the SNPs rs2304256 in exon 8 and rs12720356 in exon 15 of *Tyk2.*

Six of the SNPs were genotyped by multiplex minisequencing (fluorescence single-base extension) with the SNPstream Genotyping System (Beckman Coulter, Fullerton, CA) (36), and 2 SNPs (rs752637 and rs2304256) were genotyped using a homogeneous single-base extension assay with fluorescence polarization detection (FP-TDI; PerkinElmer, Emeryville, CA) (37). Four SNPs in the IRF-5 gene (rs729302, rs3757385, rs2004640, and rs3807306) were genotyped for replication in samples from the Dutch Early Arthritis Clinic study, using the SNPstream System as described above. The SNPs were genotyped at the SNP technology platform in Uppsala (www.genotyping.se). The sequences of the polymerase chain reaction and minisequencing primers used in the genotyping assays are available from the corresponding author upon request.

The mean SNP genotype call rate in the EIRA samples was 94% (range 92–97%), and the accuracy estimated from 3,700 genotype comparisons between repeated assays (18% of the genotypes) was 99.9%. In the Early Arthritis Clinic study samples, the mean call rate was 98% (range 95–99%), and the accuracy based on repeated determination of 10% of the genotypes was 100%. All SNPs exhibited Hardy-Weinberg equilibrium (P > 0.01 by Fisher's exact test) in both sample sets.

Statistical Analysis

Unadjusted 2-tailed P values calculated by Fisher's exact test were used to compare the SNP allele frequencies in the RA patient groups versus the controls and for assessing Hardy-Weinberg equilibrium of the SNP genotypes. Haploview, version 3.2. (38) was used to construct the haplotypes, to calculate pairwise D' and r^2 values for linkage disequilibrium, to perform haplotype association analysis using a chi-square test, and to perform permutation tests. Odds ratios (ORs) were calculated using the formula [i(G1)/i(G2)]/[j(G1)/j(G2)], where i and j are the allele counts in patients with RA and in controls, respectively.

Prior to performing a pooled analysis of the 2 independent cohorts, lack of heterogeneity in the populations was confirmed using the Breslow-Day test of homogeneity (39). The EasyMA 2001 free software package (39) was used for the pooled analysis, using the Mantel-Haenszel test for calculating ORs. The EIRA and Early Arthritis Clinic cohorts demonstrated P values for homogeneity > 0.01 for all 4 loci tested. The frequencies of ANAs in the EIRA patient and control sera were compared using a chi-square test.

Results

Association analysis using the genotype data for the individual SNPs from all Swedish patients and controls revealed that 4 of the 5 analyzed SNPs in the IRF-5 gene were associated with RA (P < 0.05), with SNP rs3807306 exhibiting the strongest association (P = 0.00063) (Table 1). In contrast, we did not find evidence of an association between RA and any of the SNPs in the Tyk-2 gene (P > 0.05). We also analyzed the genotype data from RA patients stratified into subgroups according to presence or absence of RF and stratified into subgroups according to presence or absence of anti-CCP autoantibodies, compared with the data from all controls. In the RF-positive patients, only SNP rs3807306 showed evidence of association with RA, whereas in the RF-negative patients 4 of the *IRF5* SNPs showed evidence of association.

No significant differences in the minor allele frequencies of any of the SNPs were observed when RF-positive and RF-negative patients were compared (data not shown). There was a slight difference between anti-CCP–positive and anti-CCP–negative patients in the minor allele frequency of SNP rs3807306 (P = 0.036). The most striking result of this analysis was the strong evidence of association observed for the 4 *IRF5* SNPs in the group of anti-CCP–negative patients. Despite the fact that this subgroup of patients (n = 590) constituted only approximately one-third of all patients with RA, evidence of association for SNP rs3807306 was stronger in this subgroup (P = 0.000091) than it was in the original analysis of all patients and controls (Table 1). This association was also evident in a codominant model (P = 0.0003; data not shown).

When the RA patients were stratified into subgroups according to a combination of presence or absence of RF and anti-CCP antibodies, most of the evidence of association was found in the group of patients who were negative for both anti-CCP and RF. In this group, which included only approximately one fifth of all patients (n = 366), the observed *P* values were 0.0035, 0.0082, and 0.0026 for SNPs rs3757385, rs2004640, and rs3807306, respectively. Stratification of the patients according to anti-CCP and RF status did not reveal any evidence of association of the 3 *Tyk2* SNPs with RA.

We did not detect any significant sex-specific differences in *Tyk2* or *IRF5* SNP allele frequencies in the RA patient or control groups (data not shown). The minor allele frequencies of the *IRF5* SNPs were, however, lower in the male patients who were negative for both anti-CCP and RF (n = 100) than in the male controls (n = 250), and a significantly lower minor allele frequency was noted for SNP no. rs2004640 in this subgroup compared with that in male controls (0.35 versus 0.49; *P* = 0.00091). Table 1 shows the ORs for the 5 *IRF5* SNPs in the subgroups of anti-CCP–positive and anti-CCP–negative patients. The frequency of the minor, variant alleles of the associated SNPs was lower in the RA patient groups than in the controls, reducing the risk of RA to 0.7–0.8 compared with carriers of the major alleles. Thus, these alleles appeared to be protective against RA.

Table 1. Analysis of the association of individual SNPs in the IRF-5 and Tyk-2 genes with RA in all Swedish patients and controls, and in Swedish patients after stratification based on the presence or absence of RF or anti-CCP antibodies*

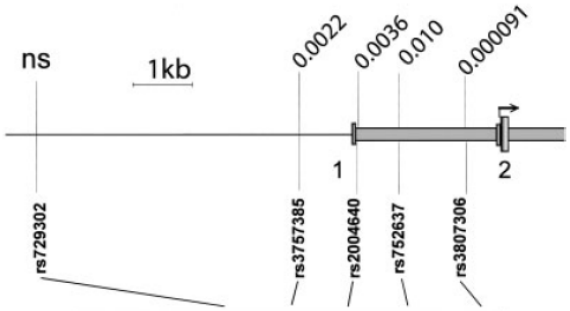
	RefSNP no., major/minor SNP allele							
	rs729302, A/C	rs3757385, C/A	rs2004640, T/G	rs752637, G/A	rs3807306, A/C	rs12720356, T/G	rs91755, G/T	rs2304256, C/A
Minor allele frequency								
Controls (n = 861)	0.31	0.36	0.48	0.37	0.50	0.09	0.48	0.30
All patients (n = 1,530)	0.32	0.31	0.44	0.33	0.45	0.09	0.47	0.29
RF+ patients (n = 956)	0.33	0.33	0.45	0.34	0.46	0.08	0.47	0.28
RF− patients (n = 489)	0.31	0.30	0.42	0.32	0.44	0.11	0.46	0.29
Anti-CCP+ patients (n = 919)	0.33	0.32	0.45	0.34	0.47	0.08	0.47	0.29
Anti-CCP− patients (n = 590)	0.31	0.30	0.43	0.32	0.43	0.11	0.47	0.28
P versus controls†								
All patients (n = 1,530)	NS	0.0012	0.0067	0.007	0.00063	NS	NS	NS
RF+ patients (n = 956)	NS	NS	NS	NS	0.008	NS	NS	NS
RF− patients (n = 489)	NS	0.0032	0.0035	0.012	0.0023	NS	NS	NS
Anti-CCP+ patients (n = 919)	NS	0.013	NS	0.043	0.035	NS	NS	NS
Anti-CCP− patients (n = 590)	NS	0.0022	0.0036	0.01	0.000091	NS	NS	NS
OR (95% CI)‡								
Anti-CCP+ patients (n = 919)	1.09 (0.94–1.26)	0.83 (0.72–0.96)	0.89 (0.77–1.01)	0.86 (0.75–0.99)	0.86 (0.75–0.98)	ND	ND	ND
Anti-CCP− patients (n = 590)	1.02 (0.86–1.19)	0.77 (0.65–0.91)	0.80 (0.68–0.92)	0.81 (0.69–0.95)	0.73 (0.62–0.85)	ND	ND	ND

* Single-nucleotide polymorphisms (SNPs) rs729302, rs3757385, rs2004640, rs752637, and rs3807306 are in the interferon regulatory factor 5 (IRF-5) gene; SNPs rs12720356, rs91755, and rs2304256 are in the Tyk-2 gene. Rheumatoid factor (RF) status could not be determined in 85 patients, and anti–cyclic citrullinated peptide antibody (anti-CCP) status could not be determined in 21 patients. 95% CI = 95% confidence interval; NS = not significant; ND = not determined.

† Unadjusted 2-tailed *P* values comparing each patient group with controls were calculated using Fisher's exact test. The empirical *P* values for association with anti-CCP–negative rheumatoid arthritis (RA) (100,000 permutations obtained using Haploview software, version 3.2) were *P* = 0.024 for SNP rs3757385, *P* = 0.042 for SNP rs2004640, and *P* = 0.0008 for SNP rs3807306. The *P* value for SNP rs752637 was not significant.

‡ Odds ratios (ORs) were calculated using the formula $[i(G_1)/i(G_2)]/[j(G_1)/j(G_2)]$, where *i* and *j* are allele counts in patients with RA and in controls, respectively.

Figure 1 shows the positions of the SNPs in the IRF-5 gene and the pairwise linkage disequilibrium values between the SNPs rs3757385, rs2004640, rs752637, and rs3807306 in the control samples from Sweden. Table 2 shows the frequencies of the haplotypes formed by these 4 IRF-5 SNPs in anti-CCP–negative RA patients, anti-CCP–positive RA patients, and controls. The 3 most common haplotypes, H1–H3, accounted for almost 90% of the genetic variation conferred by these SNPs in the sample set. The most common haplotype, H1, which is formed by the more frequent alleles of the SNPs, appeared to be a risk haplotype for anti-CCP–negative RA, with an OR of 1.23 (95% confidence interval [95% CI] 1.06–1.43). The frequency of the second most common haplotype, H2, which is formed by the minor alleles of the 4 SNPs, displayed the strongest evidence of association with RA (P = 0.000080), and was associated with protection against anti-CCP–negative RA, with an OR of 0.71 (95% CI 0.60–0.84). For anti-CCP–positive RA, the OR for haplotype H2 was 0.85 (95% CI 0.74–0.98).



rs729302		0.07	0.35	0.07	0.18
rs3757385	0.30		0.63	0.99	0.38
rs2004640	0.85	1.0		0.62	0.67
rs752637	0.30	1.0	0.99		0.38
rs3807306	0.64	0.81	0.85	0.81	

D` r^2

Figure 1. Location and pairwise linkage disequilibrium values of the single-nucleotide polymorphisms (SNPs) in the interferon regulatory factor 5 (IRF-5) gene. Top, Location of the SNPs in the IRF-5 promoter and first intron. P values for association with rheumatoid arthritis in Swedish patients who were negative for anti–cyclic citrullinated peptide antibodies are shown at the top. Exons 1 and 2 are indicated. NS _ not significant.
Bottom, Pairwise linkage disequilibrium values for the SNPs in the Swedish control samples.

Table 2. Analysis of the association of IRF-5 haplotypes with RA in Swedish patients*

	Haplotype				
	H1 (CTGA)	H2 (AGAC)	H3 (CGGC)	H4 (CTGC)	H5 (AGAA)
Haplotype frequency					
Controls	0.46	0.33	0.11	0.06	0.03
Anti-CCP+ patients	0.49	0.30	0.12	0.05	0.03
Anti-CCP− patients	0.51	0.26	0.11	0.05	0.05
P for association					
Anti-CCP+ patients	0.088	0.026	0.51	0.65	0.33
Anti-CCP− patients	0.006	0.000069	0.73	0.99	0.049
OR (95% CI)					
Anti-CCP+ patients	1.12 (0.98–1.283)	0.85 (0.74–0.98)	1.07 (0.87–1.32)	0.93 (0.69–1.25)	0.82 (0.56–1.21)
Anti-CCP− patients	1.23 (1.06–1.43)	0.71 (0.60–0.84)	0.96 (0.75–1.21)	1.00 (0.71–1.38)	1.45 (0.99–2.09)

* Haplotypes were formed by the 4 IRF-5 SNPs rs3757385, rs2004640, rs752637, and rs3807306. The empirical *P* value for association between the AGAC haplotype and RA in anti-CCP–negative patients (100,000 permutations obtained using Haploview software, version 3.2) was _0.001. See Table 1 for definitions.

To attempt to replicate the association between the IRF-5 gene and anti-CCP–negative RA, we genotyped 4 of the IRF-5 SNPs in an unrelated cohort of Dutch RA patients (Table 3). The SNP allele frequencies were similar in the Swedish and Dutch patient and control groups, and evidence of association was observed for SNP rs2004640 in the group of anti-CCP– negative patients (P = 0.024). Also, P values for SNPs rs729302 and rs3757385 (P = 0.072 and P = 0.091, respectively) suggested an association in the anti-CCP–negative patients, but, possibly due to the small number of individuals in this group, the evidence of this association did not reach the formal threshold for statistical significance.

In a combined analysis of all 741 anti-CCP–negative patients in the cohorts from both populations, which were genetically homogeneous (P > 0.01 by heterogeneity test), strong evidence of association of SNPs rs2004640 and rs3807306 with anti-CCP–negative RA was exhibited (P < 0.001). Table 3 shows the Ors for the 4 IRF-5 SNPs analyzed in the Dutch cohort and the combined ORs for all anti-CCP–negative patients in both cohorts. In the Dutch samples the haplotypes formed by SNPs rs3757385, rs2004640, and rs3807306 occurred with similar frequencies as in the Swedish samples (data not shown).

Table 3. Analysis of association of SNPs in the IRF-5 gene with RA in all Dutch patients and controls, and in Dutch and Swedish patients after stratification based on the presence or absence of anti-CCP antibodies*

	RefSNP no., major/minor SNP allele			
	rs729302, A/C	rs3757385, C/A	rs2004640, T/G	rs3807306, A/C
Minor allele frequency				
Dutch controls (n = 181)	0.35	0.40	0.51	0.49
All Dutch patients (n = 387)	0.30	0.36	0.45	0.48
Anti-CCP− Dutch patients (n = 151)	0.28	0.33	0.42	0.46
P versus controls†				
All Dutch patients (n = 387)	0.17	0.27	0.08	0.41
Anti-CCP− Dutch patients (n = 151)	0.072	0.091	0.024	0.279
Anti-CCP− Dutch and Swedish patients (n = 741)‡	0.23	0.0019	0.00089	0.00029
OR (95% CI)				
All Dutch patients (n = 387)	0.83 (0.62–1.09)	0.80 (0.66–1.13)	0.80 (0.61–1.04)	0.90 (0.69–1.17)
Anti-CCP− Dutch patients (n = 151)	0.73 (0.51–1.04)	0.76 (0.54–1.06)	0.70 (0.50–0.96)	0.84 (0.61–1.16)
Anti-CCP− Dutch and Swedish patients (n = 741)§	0.98 (0.85–1.13)	0.81 (0.70–0.92)	0.75 (0.65–0.86)	0.78 (0.68–0.89)

* SNPs rs729302, rs3757385, rs2004640, and rs3807306 are in the IRF-5 gene. See Table 1 for definitions.
† Unadjusted 2-tailed *P* values comparing patient groups with controls were calculated using Fisher's exact test.
‡ Fisher's procedure was used for combining the *P* values for anti-CCP–negative RA in the Dutch and Swedish data sets. *P* values for the Swedish anti-CCP–negative patients are shown in Table 1.
§ EasyMA 2001 software was used for the pooled analysis, using the Mantel-Haenszel test to calculate ORs.

The strongest evidence of association with RA in the Swedish cohort was observed for the *IRF5* SNP rs3807306. This SNP is in relatively high linkage disequilibrium with SNP rs2004640 (Figure 1), which is associated with SLE, and the minor alleles of these 2 SNPs are on the same protective haplotype in both SLE and RA. To test for phenotypic similarities between SLE and anti-CCP–negative RA, we measured ANAs, which are characteristic of SLE, in a subgroup of 385 RA patients and 100 age- and sex-matched controls from the EIRA cohort. We noted an increased frequency of ANAs in the RA patient samples (78 of 385) compared with the healthy control sera (5 of 100) ($P = 0.0005$), but there was no difference between anti-CCP–positive and anti-CCP–negative RA patients with respect to the frequency or staining intensity of ANAs (data not shown).

Discussion

The data presented here provide strong evidence of an association between IRF-5 gene variants and RA, and show that the IFN pathway can be involved in this disease. Interestingly, in both the Swedish and Dutch cohorts we detected the most significant association with the *IRF5* SNPs in the subgroup of anti-CCP–negative RA patients, which constituted approximately one-third of the RA patient population. Previously, no gene, with the exception of the HLA–DR3 allele (17,40), has been found to be exclusively associated with anti-CCP–negative RA. This is particularly important because the pathogenesis of anti-CCP–negative disease is poorly understood, and specific drugs for this subset of patients with RA are lacking.

Other genes that have displayed an association with RA, such as the shared epitope alleles (17) and the *PTPN22* gene (18), are associated with anti-CCP–positive RA. Anti-CCP–positive and anti-CCP–negative RA patients have clinically similar phenotypes at diagnosis, but can differ with respect to the course of the disease, and anti-CCP–negative RA patients usually develop a milder form of the disease (41). The results of the present study further support the notion that genetic background is important in the differences observed between anti-CCP–positive and anti-CCP–negative RA.

The relevant SNPs are located in or close to the first intron of the IRF-5 gene. This region contains a CpG island, which is an indication of regulatory activity, and several promoters, including one with an IFNstimulated response element site that is activated by type I IFN and one with an IRF element site that is activated by IRF-5 and IRF-1 (42). SNP rs2004640 is located in a 5' splice donor site of an alternate exon 1, and recent data indicate that there is an allele-specific effect of this SNP on the expression pattern of different isoforms of the IRF-5 gene (8).

Because IRF-5 is an important mediator of signals from TLR-7–TLR-9 (29,43), we speculate that certain expression patterns of the IRF-5 gene can modify IRF-5–dependent induction of type I IFN, proinflammatory cytokines (e.g., IL-6, TNFα, IL-12, and IL-1β), and several chemokines (44,45). Furthermore, the level of TLR-7 expression is increased in RA synovium and can contribute to synergistic cytokine production by dendritic cells (46). These cytokines can obviously influence the development and expression of inflammatory diseases, both at the level of the underlying autoimmune process and by promoting the inflammatory process. However, IRF-5 can also increase the expression levels of several genes coding for proteins that mediate cell growth arrest and apoptosis (45). Consequently, polymorphisms within the IRF-5 gene may affect several cellular functions of importance for RA susceptibility, besides the type I IFNs. However, understanding the possible functions of the allele variants of the IRF-5 gene on the molecular level requires further experimental studies.

The clear association of SNPs in the IRF-5 gene with anti-CCP–negative RA that we observed in the present study leads us to believe that the pathogenesis of this variant form of RA may be similar to the proposed pathogenesis of SLE that we described earlier (2). The findings of the present study suggest that future investigations of anti-CCP–negative RA should focus on the role of IFNα, as well as on the role of environmental agents, such as infections, that may trigger IFNα-related pathways. In a more general sense, the present clear demonstration of an association of an important non-MHC gene with anti-CCP–negative RA, but not with anti-CCP–positive RA, provides further evidence that appropriate phenotyping of RA patients and inclusion of sufficiently large study cohorts are fundamental prerequisites for future genetic and functional studies.

In summary, we now have evidence that in RA one set of genes and environmental triggers related to T cell activation and immune reactions against specific types of autoantigens (citrullinated proteins) is related to one subset of the disease, the anti-CCP–positive variant, whereas other sets of genes appear to be related to the other major subset of the disease, the anti-CCP–negative one. Thus, anti-CCP–positive RA and anti-CCP–negative RA should be considered as phenotypically distinct subsets of RA with respect to their genetic and environmental risk factors, as well as with respect to their molecular pathogenesis. In fact, we suggest that these subsets of RA should be regarded as 2 distinct diseases and treated as such in future studies.

Addendum.
After submission of this manuscript, reports of 2 studies showing no association between RA and 2 SNPs in the IRF-5 gene were published (47,48). The reasons for this seeming discrepancy could be related to the fact that the sizes of the cohorts used in the present study were larger and, perhaps more importantly, that patients in the present study were stratified according to anti-CCP antibody status.

References

(1) Theofilopoulos AN, Baccala R, Beutler B, Kono DH. Type I interferons (α/β) in immunity and autoimmunity. Annu Rev Immunol 2005;23:307–36.

(2) Ronnblom L, Eloranta ML, Alm GV. The type I interferon system in systemic lupus erythematosus [review]. Arthritis Rheum 2006; 54:408–20.

(3) Bave U, Nordmark G, Lovgren T, Ronnelid J, Cajander S, Eloranta ML, et al. Activation of the type I interferon system in primary Sjo¨gren's syndrome: a possible etiopathogenic mechanism. Arthritis Rheum 2005;52:1185–95.

(4) Nestle FO, Conrad C, Tun-Kyi A, Homey B, Gombert M, Boyman O, et al. Plasmacytoid predendritic cells initiate psoriasis through interferon-α production. J Exp Med 2005;202:135–43.

(5) Greenberg SA, Pinkus JL, Pinkus GS, Burleson T, Sanoudou D, Tawil R, et al. Interferon-α/β-mediated innate immune mechanisms in dermatomyositis. Ann Neurol 2005;57:664–78.

(6) Stewart TA. Neutralizing interferon = as a therapeutic approach to autoimmune diseases. Cytokine Growth Factor Rev 2003;14: 139–54.

(7) Sigurdsson S, Nordmark G, Goring HH, Lindroos K, Wiman AC, Sturfelt G, et al. Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. Am J Hum Genet 2005;76:528–37.

(8) Graham RR, Kozyrev SV, Baechler EC, Reddy MV, Plenge RM, Bauer JW, et al. A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. Nat Genet 2006;38:550–5.

(9) Gota C, Calabrese L. Induction of clinical autoimmune disease by therapeutic interferon-=. Autoimmunity 2003;36:511–8.

(10) Klareskog L, Stolt P, Lundberg K, Kallberg H, Bengtsson C, Grunewald J, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA–DR (shared epitope)–restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006;54:38–46.

(11) Van Zeben D, Hazes JM, Zwinderman AH, Cats A, van der Voort EA, Breedveld FC. Clinical significance of rheumatoid factors in early rheumatoid arthritis: results of a follow up study. Ann Rheum Dis 1992;51:1029–35.

(12) Renaudineau Y, Jamin C, Saraux A, Youinou P. Rheumatoid factor on a daily basis. Autoimmunity 2005;38:11–6.

(13) Schellekens GA, Visser H, de Jong BA, van den Hoogen FH, Hazes JM, Breedveld FC, et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum 2000;43:155–63.

(14) Rantapaa-Dahlqvist S, de Jong BA, Berglin E, Hallmans G, Wadell G, Stenlund H, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003;48:2741–9.

(15) Ronnelid J, Wick MC, Lampa J, Lindblad S, Nordmark B, Klareskog L, et al. Longitudinal analysis of citrullinated protein/peptide antibodies (anti-CP) during 5 year follow up in early rheumatoid arthritis: anti-CP status predicts worse disease activity and greater radiological progression. Ann Rheum Dis 2005;64: 1744–9.

(16) Padyukov L, Silva C, Stolt P, Alfredsson L, Klareskog L, for the Epidemiological Investigation of Rheumatoid Arthritis Study Group. A gene–environment interaction between smoking and shared epitope genes in HLA–DR provides a high risk of seropositive rheumatoid arthritis. Arthritis Rheum 2004;50:3085–92.

(17) Irigoyen P, Lee AT, Wener MH, Li W, Kern M, Batliwalla F, et al. Regulation of anti–cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA–DR3 and the shared epitope alleles. Arthritis Rheum 2005;52:3813–8.

(18) Plenge RM, Padyukov L, Remmers EF, Purcell S, Lee AT, Karlson EW, et al. Replication of putative candidate-gene associations with rheumatoid arthritis in >4,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. Am J Hum Genet 2005;77:1044–60.

(19) Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004;75:330–7.

(20) Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, et al. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat Genet 2003;34:395–402.

(21) Rodriguez MR, Nunez-Roldan A, Aguilar F, Valenzuela A, Garcia A, Gonzalez-Escribano MF. Association of the CTLA4 3' untranslated region polymorphism with the susceptibility to rheumatoid arthritis. Hum Immunol 2002;63:76–81.

(22) Wesoly J, van der Helm-van Mil AH, Toes RE, Chokkalingam AP, Carlton VE, Begovich AB, et al. Association of the PTPN22 C1858T single-nucleotide polymorphism with rheumatoid arthritis phenotypes in an inception cohort. Arthritis Rheum 2005;52: 2948–50.

(23) Van Holten J, Smeets TJ, Blankert P, Tak PP. Expression of interferon β in synovial tissue from patients with rheumatoid arthritis: comparison with patients with osteoarthritis and reactive arthritis. Ann Rheum Dis 2005;64:1780–2.

(24) Lande R, Giacomini E, Serafini B, Rosicarelli B, Sebastiani GD, Minisola G, et al. Characterization and recruitment of plasmacytoid dendritic cells in synovial fluid and tissue of patients with chronic inflammatory arthritis. J Immunol 2004;173:2815-24.

(25) Jongbloed SL, Lebre MC, Fraser AR, Gracie JA, Sturrock RD, Tak PP, et al. Enumeration and phenotypical analysis of distinct dendritic cell subsets in psoriatic arthritis and rheumatoid arthritis. Arthritis Res Ther 2006;8:R15.

(26) Cavanagh LL, Boyce A, Smith L, Padmanabha J, Filgueira L, Pietschmann P, et al. Rheumatoid arthritis synovium contains plasmacytoid dendritic cells. Arthritis Res Ther 2005;7:R230–40.

(27) Firestein GS. Immunologic mechanisms in the pathogenesis of rheumatoid arthritis. J Clin Rheumatol 2005;11 Suppl 3:S39–44. THE IRF-5 GENE IN RA 2209

(28) Shaw MH, Boyartchuk V, Wong S, Karaghiosoff M, Ragimbeau J, Pellegrini S, et al. A natural mutation in the Tyk2 pseudokinase domain underlies altered susceptibility of B10.Q/J mice to infection and autoimmunity. Proc Natl Acad Sci U S A 2003;100: 11594–9.

(29) Schoenemeyer A, Barnes BJ, Mancl ME, Latz E, Goutagny N, Pitha PM, et al. The interferon regulatory factor, IRF5, is a central mediator of Toll-like receptor 7 signaling. J Biol Chem 2005;280: 17005–12.

(30) Takaoka A, Yanai H, Kondo S, Duncan G, Negishi H, Mizutani T, et al. Integral role of IRF-5 in the gene induction programme activated by Toll-like receptors. Nature 2005;434:243–9.

(31) Izaguirre A, Barnes BJ, Amrute S, Yeow WS, Megjugorac N, Dai J, et al. Comparative analysis of IRF and IFN-α expression in human plasmacytoid and monocyte-derived dendritic cells. J Leukoc Biol 2003;74:1125–38.

(32) Barnes BJ, Moore PA, Pitha PM. Virus-specific activation of a novel interferon regulatory factor, IRF-5, results in the induction of distinct interferon α genes. J Biol Chem 2001;276:23382–90.

(33) Arnett FC, Edworthy SM, Bloch DA, McShane DJ, Fries JF, Cooper NS, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.

(34) Van Aken J, van Bilsen JH, Allaart CF, Huizinga TW, Breedveld FC. The Leiden Early Arthritis Clinic. Clin Exp Rheumatol 2003;21(5 Suppl 31):S100–5.

(35) Van der Horst-Bruinsma IE, Speyer I, Visser H, Breedveld FC, Hazes JM. Diagnosis and course of early-onset arthritis: results of a special early arthritis clinic compared to routine patient care. Br J Rheumatol 1998;37:1084–8.

(36) Bell PA, Chaturvedi S, Gelfand CA, Huang CY, Kochersperger M, Kopla R, et al. SNPstream UHT: ultra-high throughput SNP genotyping for pharmacogenomics and drug discovery [published erratum appears in Biotechniques 2003;34:496]. Biotechniques 2002;Suppl:70–2, 74, 76–7.

(37) Hsu TM, Chen X, Duan S, Miller RD, Kwok PY. Universal SNP genotyping assay with fluorescence polarization detection. Biotechniques 2001;31:560–8.

(38) Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 2005;21: 263–5.

(39) Cucherat M, Boissel JP, Leizorovicz A, Haugh MC. EasyMA: a program for the meta-analysis of clinical trials. Comput Methods Programs Biomed 1997;53:187–90.

(40) Verpoort KN, van Gaalen FA, van der Helm-van Mil AH, Schreuder GM, Breedveld FC, Huizinga TW, et al. Association of HLA–DR3 with anti–cyclic citrullinated peptide antibody–negative rheumatoid arthritis. Arthritis Rheum 2005;52:3058–62.

(41) Van der Helm-van Mil AH, Verpoort KN, Breedveld FC, Toes RE, Huizinga TW. Antibodies to citrullinated proteins and differences in clinical progression of rheumatoid arthritis. Arthritis Res Ther 2005;7:R949–58.

(42) Mancl ME, Hu G, Sangster-Guity N, Olshalsky SL, Hoops K, Fitzgerald-Bocarsly P, et al. Two discrete promoters regulate the alternatively spliced human interferon regulatory factor-5 isoforms: multiple isoforms with distinct cell type-specific expression, localization, regulation, and function. J Biol Chem 2005;280:21078–90.

(43) Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell 2006;124:783–801.

(44) Barnes BJ, Kellum MJ, Field AE, Pitha PM. Multiple regulatory domains of IRF-5 control activation, cellular localization, and induction of chemokines that mediate recruitment of T lymphocytes. Mol Cell Biol 2002;22:5721–40.

(45) Barnes BJ, Richards J, Mancl M, Hanash S, Beretta L, Pitha PM. Global and distinct targets of IRF-5 and IRF-7 during innate response to viral infection. J Biol Chem 2004;279:45194–207.

(46) Roelofs MF, Joosten LA, Abdollahi-Roodsaz S, van Lieshout AW, Sprong T, van den Hoogen FH, et al. The expression of Toll-like receptors 3 and 7 in rheumatoid arthritis synovium is increased and costimulation of Toll-like receptors 3, 4, and 7/8 results in synergistic cytokine production by dendritic cells. Arthritis Rheum 2005;52:2313–22. 47. Rueda B, Reddy MV, Gonzalez-Gay MA, Balsa A, Pascual- Salcedo D, Petersson IF, et al. Analysis of IRF5 gene functional polymorphisms in rheumatoid arthritis. Arthritis Rheum 2006;54: 3815–9.

(47) Garnier S, Dieude P, Michou L, Barbet S, Tan A, Lasbleiz S, et al. The systemic lupus erythematosus new genetic factor IRF5 rs2004640-T allele is not linked to, nor associated with rheumatoid arthritis, in a family-based study from the French Caucasian population. Ann Rheum Dis 2006. E-pub ahead of print.

Chapter 15

Common variants at CD40 and other loci confer risk of rheumatoid arthritis

S. Raychaudhuri, E.F. Remmers, A.T. Lee, R. Hackett, C. Guiducci, N.P. Burtt, L. Gianniny, B.D. Korman, L. Padyukov, F.A.S. Kurreeman, M. Chang, J.J. Catanese, B. Ding, S. Wong, A.H.M. van der Helm-van Mil, B.M. Neale, J. Coblyn, J. Cui, P.P. Tak, G.J. Wolbink, J Bart A Crusius, I.E. van der Horst-Bruinsma, L.A. Criswell, C.I. Amos, M.F. Seldin, N.A. Shadick, M.E. Weinblatt, N. de Vries, J. Worthington, M. Seielstad, D.L. Kastner, K.G. Ardlie, L. Alfredsson, K.H. Costenbader, D. Altshuler, T.W.J. Huizinga, R.E.M. Toes, E.W. Karlson, A.B. Begovich, L. Klareskog, P.K. Gregersen, M.J. Daly, R.M. Plenge

Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts 02142, USA.
Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA.
Center for Human Genetic Research and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA.
Genetics and Genomics Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, US National Institutes of Health, Bethesda, Maryland 20892, USA.
The Feinstein Institute for Medical Research, North Shore-Long Island Jewish Health System, Manhasset, New York 11030, USA.
Rheumatology Unit, Department of Medicine, Karolinska Institutet at Karolinska University Hospital Solna, Stockholm 171 76, Sweden.
Department of Rheumatology, Leiden University Medical Centre, Leiden 2333ZA, The Netherlands.
Celera, Alameda, California 94502, USA.
Institute of Environmental Medicine, Karolinska Institutet, Stockholm 171 77, Sweden.
Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College, London SE5 8AF, UK.
Clinical Immunology and Rheumatology, Academic Medical Center, University of Amsterdam, Amsterdam NL326, The Netherlands.
Jan van Breemen Institute, Amsterdam 1056 AB, The Netherlands.
Sanquin Research Landsteiner Laboratory, Academic Medical Center, University of Amsterdam, Amsterdam 1006 AD, The Netherlands.
Laboratory of Immunogenetics, Department of Pathology and
Department of Rheumatology, VU University Medical Center, Amsterdam 1007 MB, The Netherlands.
Rosalind Russell Medical Research Center for Arthritis, Department of Medicine, University of California, San Francisco, California 94143, USA.
University of Texas M.D. Anderson Cancer Center, Houston, Texas 77030, USA.
Rowe Program in Genetics, University of California at Davis, Davis, California 95616, USA.
SeraCare Life Science, Cambridge, Massachusetts 02139, USA.
Arthritis Research Campaign (arc)–Epidemiology Unit, Stopford Building, The University of Manchester, Manchester M13 9PT, UK.
Genome Institute of Singapore, Singapore 138672.

Nat Genet. 40, 1216-1223 (2008)

Abstract

To identify rheumatoid arthritis risk loci in European populations, we conducted a meta-analysis of two published genome-wide association (GWA) studies totaling 3,393 cases and 12,462 controls[1,2]. We genotyped 31 top-ranked SNPs not previously associated with rheumatoid arthritis in an independent replication of 3,929 autoantibody-positive rheumatoid arthritis cases and 5,807 matched controls from eight separate collections. We identified a common variant at the CD40 gene locus (rs4810485, P = 0.0032 replication, P = 8.2 x 10^{-9} overall, OR = 0.87). Along with other associations near TRAF1 (refs. 2,3) and TNFAIP3 (refs. 4,5), this implies a central role for the CD40 signaling pathway in rheumatoid arthritis pathogenesis. We also identified association at the CCL21 gene locus (rs2812378, P = 0.00097 replication, P = 2.8 x 10^{-7} overall), a gene involved in lymphocyte trafficking. Finally, we identified evidence of association at four additional gene loci: MMEL1-TNFRSF14 (rs3890745, P = 0.0035 replication, P = 1.1 x 10^{-7} overall), CDK6 (rs42041, P = 0.010 replication, P = 4.0 x 10^{-6} overall), PRKCQ (rs4750316, P = 0.0078 replication, P = 4.4 x 10^{-6} overall), and KIF5A-PIP4K2C (rs1678542, P = 0.0026 replication, P = 8.8 x 10^{-8} overall).

Introduction

Rheumatoid arthritis is a systemic autoimmune disease with intraarticular inflammation as a dominant feature that affects up to 1% of the population. It can be subdivided clinically by the presence or absence of autoantibodies (antibodies to cyclic citrullinated peptide (CCP) or rheumatoid factor (RF), both of which are highly correlated to each other). Previous genetic studies have identified and validated five risk loci for autoantibody-positive RA: multiple alleles within the MHC region (6); a missense allele in the PTPN22 gene (7); two alleles at the 6q23 locus near the TNFAIP3 gene (4,5); and single alleles in the STAT4 locus (8) and TRAF1-C5 loci (2). Additional alleles at 4q27 (ref. 9), CTLA4 (ref. 10) and PADI4 (ref. 11) have suggestive associations, but have not yet been widely replicated in individuals of European ancestry.

To identify a collection of unbiased candidate rheumatoid arthritis risk loci for further investigation, we carried out a meta-analysis of SNP data from three case-control collections of European individuals from two published GWA studies1,2 (Table 1, see Methods for details). We investigated a common set of ~340,000 SNPs genotyped by the Wellcome Trust Case Control Consortium (WTCCC) with an Affymetrix 500K platform that (i) passed strict quality control criteria and (ii) were also present in the Phase II HapMap.

Table 1 Sample collections

	Case collection	Control collection	Origin	Antibody status	Cases	Controls	Genotyping platform	Case-control stratification
Meta-analysis	North American Rheumatoid Arthritis Consortium (NARAC)	New York Cancer Project, New York City	North America	100% CCP+	873	1,196	Illumina 550K	Identity-by-state clustering
3,393 cases; 12,460 controls	Epidemiological Investigation of Rheumatoid Arthritis (EIRA)	EIRA	Sweden	100% CCP+	660	658	Illumina 317K	Epidemiologically matched, identity-by-state clustering
	Wellcome Trust Case Control Consortium (WTCCC)	Shared controls, multiple non-autoimmune diseases	United Kingdom	80% CCP+, 84% RF+	1,860	10,606	Affymetrix 500K	Geographically matched
Stage 1 replication	Nurses Health Study (NHS)	NHS	North America	100% RF+ or CCP+	257	411	Sequenom iPlex	Epidemiologically matched
	Brigham Rheumatoid Arthritis Sequential Study (BRASS)	National Institutes of Mental Health (NIMH)	Boston, USA	100% CCP+	407	814	Sequenom iPlex, Affymetrix 500K	Case-control matching with GWAS data
1,089 cases; 1,862 controls	NARAC II	New York Cancer Project, New York City	North America	100% CCP+	425	637	Sequenom iPlex	Case-control matching with ancestry informative markers
	NARAC III	Publicly available shared controls	North America	100% CCP+	869	1,303	Illumina 317K	Case-control matching with GWAS data
Stage 2 replication	Genomics Collaborative Initiative (GCI)	GCI	North America	100% RF+	457	460	Kinetic PCR	Epidemiologically matched
	Leiden University Medical Center (LUMC)	LUMC	Leiden, The Netherlands	100% RF+ or CCP+	528	540	Kinetic PCR	Geographically matched
2,840 cases; 3,945 controls	EIRA-II	EIRA-II	Sweden	100% CCP+	435	412	Sequenom iPlex	Epidemiologically matched
	Genetics Network Rheumatology Amsterdam (GENRA)	GENRA	Amsterdam, The Netherlands	100% CCP+	551	1,230	Sequenom iPlex	Geographically matched

GWA data from three meta-analysis collections were used to identify candidate SNPs for replication. The replication set was divided into two stages: stage 1 replication (three collections) and stage 2 replication (five collections). For each collection we list the geographic origin, the source of the controls, the autoantibody status of cases, and the number of cases and controls. We also list the genotyping technology used to type SNPs of interest. Finally, we specify the strategy used to correct for case-control population stratification.

We used the software package IMPUTE[12] to determine genotypes of these SNPs in individuals from Sweden (Epidemiological Investigation of Rheumatoid Arthritis, EIRA) and North America (North American Rheumatoid Arthritis Consortium, NARAC) on the basis of available Illumina platform SNP data (Supplementary Fig. 1 online). To conduct a meta-analysis of SNP association with rheumatoid arthritis risk, we used the Cochran-Mantel-Haenszel (CMH) statistical test using genotype counts from the WTCCC and imputed probabilistic allele dosages in EIRA and NARAC. The CMH test allowed us to conduct a stratified analysis that maintained the three case-control collections as separate strata. CMH also allowed for further sub-stratification of EIRA and NARAC individuals into more homogenous subgroups using identity-by-state similarity for SNPs across the genome[2] to correct for residual population stratification. This resulted in improved genomic control inflation for both EIRA (λ_{GC} = 1.03) and NARAC (λ_{GC} = 1.20). As there was little evidence of population stratification in the WTCCC (λ_{GC} = 1.06), we did not further substratify those individuals.

After calculating case-control CMH association statistics in the GWA meta-analysis, we observed minimal inflation for SNPs outside the MHC region (λ_{GC} = 1.09, λ_{GC} = 1.02 after normalizing to a 1,000 case and control collection, Supplementary Fig. 2 online). Thus, there was little evidence of bias due to technical artifact or population stratification. In Table 2 we present association statistics for validated and suggestive rheumatoid arthritis risk loci in European populations. Of the confirmed non-MHC risk loci, we observed association at PTPN22, 6q23/(containing TNFAIP3), STAT4 and TRAF1-C5. We also observed evidence of association at 4q27 (containing the IL2 and IL21 genes) and CTLA4, but not PADI4. The 4q27 result is an independent replication, suggesting that this is a true-positive association. CTLA4 replicates in the WTCCC (P = 0.026), providing further support for the role of this locus in rheumatoid arthritis, as suggested by previous studies in EIRA and NARAC[10].

Table 2 Meta-analysis results from regions previously associated with rheumatoid arthritis

SNP					EIRA		NARAC		WTCCC		Meta-analysis		Minor allele frequency		Meta-analysis odds ratio	Published
Prior study	Meta-analysis	r^2	Locus	Region	χ^2	P	χ^2	P	χ^2	P	χ^2	P	RA	Non-RA	(95% CI)	odds ratio
rs2240340	rs11203367	0.97	*PADI4*	1p36	3.6	0.058	0.4	0.55	0.2	0.64	0.3	0.59	0.42	0.42	1.02 (0.96–1.08)	1.10
rs2476601	rs6679677	1	*PTPN22*	1p13	13.1	0.00029	30.2	3.8×10^{-08}	151.8	7.1×10^{-35}	184.3	5.7×10^{-42}	0.17	0.10	1.79 (1.65–1.94)	1.75
rs7574865	rs11893432	0.74	*STAT4*	2q32	2.7	0.10	3.9	0.050	6.0	0.014	11.9	0.00055	0.21	0.18	1.14 (1.06–1.23)	1.32
rs3087243	rs3087243	n.a.	*CTLA4*	2q33	4.2	0.041	4.3	0.037	4.9	0.026	11.8	0.00060	0.40	0.44	0.90 (0.85–0.95)	0.88
rs6822844	rs4572894	0.47	*IL2, IL21*	4q27	1.8	0.18	4.8	0.029	1.8	0.18	6.4	0.011	0.27	0.29	0.92 (0.86–0.97)	0.72
*HLA-DRB1*04*	rs6457620	n.a.	*HLA-DRB1*	6p21	145.1	2.0×10^{-33}	321.6	6.4×10^{-72}	439.5	1.4×10^{-97}	846.8	3.6×10^{-186}	0.72	0.50	2.55 (2.40–2.71)	~3
rs6920220	rs6920220	n.a.	*OLIG3, TNFAIP3*	6q23	7.6	0.0059	7.8	0.0053	22.6	2.0×10^{-06}	36.5	1.5×10^{-09}	0.26	0.22	1.24 (1.16–1.32)	1.22
rs10499194	rs1167223	0.5	*OLIG3, TNFAIP3*	6q23	1.9	0.17	4.9	0.026	10.7	0.0011	17.0	0.000038	0.36	0.39	0.88 (0.83–0.93)	0.75
rs3761847	rs10118357	0.97	*TRAF1, C5*	9q33	7.9	0.0050	22.1	2.6×10^{-06}	0.0	0.96	9.3	0.0023	0.46	0.44	1.10 (1.04–1.16)	1.32

For each SNP, we identified the best proxy in our study and the calculated LD to that proxy on the basis of CEU HapMap (reported as r2). We present the w2 score and two-tailed P value in EIRA, NARAC and WTCCC, and in the meta-analysis.

For each of the collections we further addressed potential case-control population stratification by either (i) using epidemiologically matched samples or (ii) matching cases and controls with ancestry informative genetic data; detailed strategies for each collection are described in the Supplementary Note online.

Our replication collection consisted of eight independent casecontrol collections totaling 3,929 autoantibody (either CCP or RF) positive rheumatoid arthritis cases and 5,807 matched controls, all self-described as white and of European ancestry (Table 1). The presence of CCP or RF autoantibodies assures specificity for the diagnosis of rheumatoid arthritis and helps to minimize clinical heterogeneity across the eight collections. For each of the collections we further addressed potential case-control population stratification by either (i) using epidemiologically matched samples or (ii) matching cases and controls with ancestry informative genetic data; detailed strategies for each collection are described in the Supplementary Note online.

A total of 31 of these SNPs were successfully genotyped in all three stage 1 replication collections. The 17 most significant SNPs were genotyped in the stage 2 replication collections. For each SNP we calculated the replication P value using a one-tailed CMH statistic across the replication collections, and for those that replicated with P < 0.05, we calculated an overall P value (replication and the three meta-analysis collections) using a two-tailed CMH statistic.

Testing in our complete replication set identified rheumatoid arthritis–associated alleles (Table 3 and Supplementary Tables 2 and 3 online). In replication genotyping, we observed that 6 out of31 SNPs obtained P ≤ 0.01; this is significantly more than expected by chance alone (P = 9 x 10^{-7} by Poisson). Figure 1 illustrates the observed one-tailed CMH replication z scores, which clearly show that our results are enriched for z > 2 values (which corresponds to P < 0.023). Also, 4 of the 340,000 SNPs tested initially in the meta-analysis are associated with P < 5 x 10^{-7} in joint analysis; this is also significantly more than expected by chance alone (P = 3 x 10^{-5} by Poisson).

After excluding published risk loci (including those in the MHC region) and correcting for residual inflation by λ_{GC}, we found that 78 SNPs remained with possible associations at a P < 10^{-4} threshold (Supplementary Table 1 online). These SNPs were grouped into 38 independent regions on the basis of pairwise linkage disequilibrium (LD) estimates derived from CEU HapMap (where SNPs were grouped together if r^2 > 0.1). We tested the single most significant SNP from each region in a two-staged replication.

Our replication collection consisted of eight independent casecontrol collections totaling 3,929 autoantibody (either CCP or RF) positive rheumatoid arthritis cases and 5,807 matched controls, all self-described as white and of European ancestry (Table 1). The presence of CCP or RF autoantibodies assures specificity for the diagnosis of rheumatoid arthritis and helps to minimize clinical heterogeneity across the eight collections.

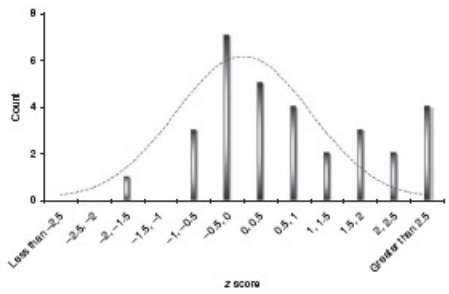


Figure 1 Enrichment of SNPs with z scores 42 in replication samples.

For each of the 31 SNPs tested, we calculated a one-sided CMH z-score statistic from our two-staged replication data. Results were calculated using either stage 1 replication samples only (n = 14 SNPs) or using both stage 1 and 2 replication samples (n = 17 SNPs). A z score of 0 corresponds to a P value of 0.5, and a z score of 1.65 corresponds to a P value of 0.05. For a random collection of unassociated SNPs, this histogram should approximate a normal distribution (dotted line).

Table 3 Newly identified SNPs associated with rheumatoid arthritis susceptibility

SNP						Meta-analysis				Replication				Joint		Breslow-Day test
ID	Chr.	Position	Locus	Allele		Frequency				Frequency						
				Major	Minor	Control	Case	*p*	OR	Control	Case	*p*	OR	*p*	OR	*p*
rs3890745	1p36	2585786	*MMEL1-TNFRSF14*	T	C	0.329	0.295	4.3×10^{-6}	0.86	0.321	0.301	0.0035	0.92	1.1×10^{-7}	0.89	0.31
rs42041	7q21	91891395	*CDK6*	C	G	0.243	0.274	5.5×10^{-5}	1.15	0.261	0.277	0.010	1.08	4.0×10^{-6}	1.11	0.43
rs2812378	9p13	34700260	*CCL21*	A	G	0.339	0.364	6.9×10^{-5}	1.13	0.334	0.355	0.00097	1.10	2.8×10^{-7}	1.12	0.67
rs4750316	10p15	6433266	*PRKCQ*	G	C	0.194	0.171	9.9×10^{-5}	0.86	0.197	0.183	0.0078	0.91	4.4×10^{-6}	0.88	0.77
rs1678542	12q13	56254982	*KIF5A-PIP4K2C*	C	G	0.373	0.352	5.4×10^{-6}	0.87	0.366	0.351	0.0026	0.92	8.8×10^{-8}	0.89	0.62
rs4810485	20q13	44181354	*CD40*	G	T	0.254	0.221	2.4×10^{-7}	0.83	0.246	0.231	0.0032	0.91	8.2×10^{-9}	0.87	0.37

We list results of 6 SNPs (out of 31 tested) that replicated with P ≤ 0.01. The first six columns list SNP characteristics. The next four columns list GWA meta-analysis results including allele frequencies, a two-tailed P value for SNP association, and an odds ratio. The next four columns list similar results for replication genotyping; significance is reported on the basis of a stratified one-tailed CMH statistic. The next two columns summarize joint (overall) analysis results. Significance is reported on the basis of a stratified two-tailed CMH statistic across all 11 sample collections. The final column lists the Breslow-Day test for heterogeneity of odds ratios across all 11 collections.

One SNP, rs4810485 in the 20q13 region, surpasses a conservative level of significance in joint analysis ($P = 8.2 \times 10^{-9}$) and thus represents a confirmed rheumatoid arthritis risk variant (Fig. 2). This SNP is located in the second intron of CD40 and is within an LD block that contains a large portion of the CD40 gene and its 5' intergenic region. A SNP in near-perfect LD with rs4810485 ($r^2 = 0.95$, rs1883832) has been associated with autoimmune thyroid disease[13], although the association has not been confirmed unequivocally[14,15]. The same allele contributes to risk in both diseases. The rs1883832 variant has been shown to influence the efficiency of CD40 protein translation by disrupting a Kozak sequence[13]. The CD40 protein is expressed on the surface of multiple immune cells, including B cells, monocytes and dendritic cells, whereas its ligand, CD154, is expressed by activated $CD4^+$ T cells. CD40-CD154 interactions play a pivotal role in provision of helper activity by $CD4^+$ T cells in immune reactions including immunoglobulin class switching, memory B-cell development and germinal center formation[16]. Null mutations in CD40 are known to cause a rare B cell–dependent hyper-IgM immunodeficiency syndrome[17].

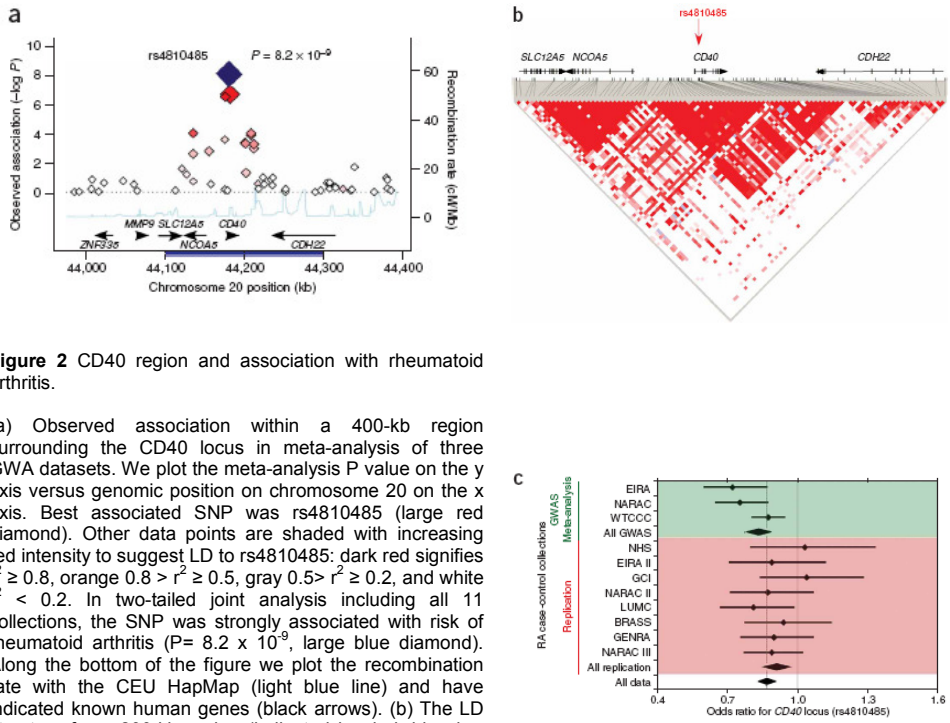


Figure 2 CD40 region and association with rheumatoid arthritis.

(a) Observed association within a 400-kb region surrounding the CD40 locus in meta-analysis of three GWA datasets. We plot the meta-analysis P value on the y axis versus genomic position on chromosome 20 on the x axis. Best associated SNP was rs4810485 (large red diamond). Other data points are shaded with increasing red intensity to suggest LD to rs4810485: dark red signifies $r^2 \geq 0.8$, orange $0.8 > r^2 \geq 0.5$, gray $0.5 > r^2 \geq 0.2$, and white $r^2 < 0.2$. In two-tailed joint analysis including all 11 collections, the SNP was strongly associated with risk of rheumatoid arthritis ($P = 8.2 \times 10^{-9}$, large blue diamond). Along the bottom of the figure we plot the recombination rate with the CEU HapMap (light blue line) and have indicated known human genes (black arrows). (b) The LD structure for a 200-kb region (indicated by dark blue bar along the x axis in a) surrounding the rs4810485 SNP (red arrow).. The rs4810485 SNP is located in the second intron of CD40. LD between SNPs is indicated by the redness of individual spots in the triangular graphic, the most intense red spots have a D'= 1. (c) Forest plot for rs4810485 association across 11 collections. For each sample collection we plot the odds ratio (small diamond) and the 95% CI. A light dashed line indicates the final odds ratio across all collections. The top three bars (green) represent GWA data and the diamond below them summarizes their meta-analyzed effect. The next eight bars (red) represent replication data and the diamond below them summarizes their meta-analyzed effect. The final diamond at the bottom of the page (clear) represents the final meta-analysis odds ratio and the 95% CI for all 11 collections.

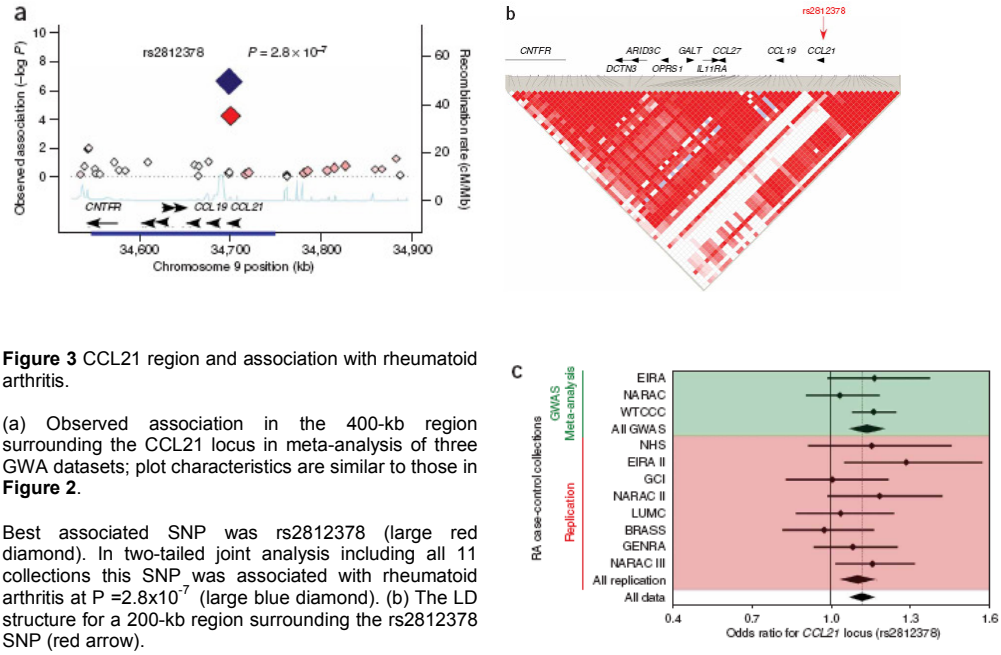


Figure 3 CCL21 region and association with rheumatoid arthritis.

(a) Observed association in the 400-kb region surrounding the CCL21 locus in meta-analysis of three GWA datasets; plot characteristics are similar to those in **Figure 2**.

Best associated SNP was rs2812378 (large red diamond). In two-tailed joint analysis including all 11 collections this SNP was associated with rheumatoid arthritis at P =2.8x10^{-7} (large blue diamond). (b) The LD structure for a 200-kb region surrounding the rs2812378 SNP (red arrow).
The rs2812378 SNP is located close to the 5' end of the CCL21. (c) Forest plot for rs2812378 association across 11 collections.

The rs2812378 SNP in the 9p13 region replicates convincingly (P = 0.00097) and has a highly suggestive level of significance with an overall P = 2.8 x 10^{-7} (Fig. 3). The SNP is located ~0.1 kb from the 5' untranslated region of the CCL21 gene, and is near a cluster of other genes including CCL19 and CCL27. However, it is in an LD block that fully includes the CCL21 coding sequence and not the other genes. The CCL21 protein is a chemokine that is involved in homing lymphocytes to secondary lymphoid organs. Expression of this chemokine is associated with ectopic lymphoid structures and has been implicated in the organization of lymphoid tissue affected by rheumatoid arthritis[18].

Most of the four other SNPs with P ≤ 0.01 in our combined stage 1 and 2 genotyping probably represent true rheumatoid arthritis susceptibility alleles, although additional validation in large sample collections is required. Of the four regions, each contains genes that are known to be critical to the immune system. The rs42041 SNP on chromosome 7 maps to a CDK6 intron; CDK6 is a ubiquitous cyclin-dependent kinase that regulates cell cycle progression, and it has been identified as a key mediator in the rapid proliferation of B cells and CD8 memory cells[19,20]. The rs1678542 SNP on chromosome 12 is ~20 kb away from the PIP4K2C locus, which has been implicated in signalling through the B-cell antigen receptor[21]. The rs3890745 SNP on chromosome 1 is ~60 kb away from TNFRSF14, which is similar to CD40 in that it is a member

of the TNF receptor super-family; it is known to bind TRAF family members including TRAF1 and is involved in activation of the transcription factors NF-κB and AP-1 (ref. 22). The rs4750316 SNP on chromosome 10 is ~100 kb away from the 3' end of the PRKCQ gene, which encodes a kinase required for the activation of the transcription factors NF-κB and AP-1, and may link the T-cell receptor signalling complex to the activation of the transcription factors[23].

A parallel UK study in this issue investigating the most significantly associated SNPs within the WTCCC study[1] provides additional evidence for associations with rs4750316 (PRKCQ), rs1678542 (KIF5A-PIP4K2C) and rs10910099 (with r^2 = 0.96 to rs3890745 in the MMEL1-TNFRSF14 locus)[24].

Population stratification is unlikely to account for these observed effects, despite the modest effect sizes observed for rheumatoid arthritis risk (0.87 ≤ OR ≤ 1.12). We were careful to control for stratification individually in each of the meta-analysis GWA studies and also in each of the eight replication collections. Furthermore, the WTCCC collection contributed the greatest number of samples to the meta-analysis, and careful investigation across 12 subregions in the UK showed little evidence of case-control stratification[1]. Each of the associations presented here is notably significant in the WTCCC alone (P < 0.001). We found no evidence that different effects were present for these six loci across the five collections with genetically matched controls and the six collections with epidemiologically matched samples (Breslow-Day P > 0.31, Table 2). Technical artifact cannot explain the associations, as all SNPs passed strict quality control criteria (Supplementary Table 4 online).

These associations provide strong evidence for the importance of the CD40 signaling pathway in autoantibody-positive rheumatoid arthritis. Our study implicates a putative functional variant that affects protein translation of the CD40 receptor. Established associations near TRAF1 and TNFAIP3 (also known as A20) already suggest the possibility that the CD40 signaling pathway mediates rheumatoid pathogenesis through NF-kB activation[25], although the rheumatoid arthritis risk variants have not yet been proven to modulate function of these genes. In particular, TRAF1 binds the CD40 receptor and cooperates with TRAF2 to activate NF-κB[26]. TRAF1 also binds TNFAIP3, which is a negative regulator of NF-κB signaling[27]. Furthermore, CD40 stimulation results in B-cell proliferation through regulation of CDK6 expression[19]. The CD40 signaling pathway has been investigated in drug development, and mouse models have demonstrated that its disruption could prevent development of immunemediated arthritis and diabetes[28,29].

In conclusion, our study has identified a rheumatoid arthritis–associated variant for European populations at the CD40 locus, provided strong evidence for association at the CCL21 locus, and also suggests association at four other loci. It also provides empirical data suggesting that additional common alleles with odds ratios ~1.15 remain to be discovered. Even under the assumption that all of these variants are true risk factors, their total percent variance explained is only 0.6% (Supplementary Note). In this study, we estimate that the total percent variance explained for all known non-MHC common genetic variants is just 3.6%. Considering that ~60% of rheumatoid arthritis risk is thought to be genetic, and one-third of this risk is from the MHC locus[30], this indicates that less than half of genetic variation can be explained by the known rheumatoid arthritis risk alleles. One possibility is that there are other non-MHC common variants that have not yet been detected. All of the variants identified in our study have very modest effects and the rheumatoid arthritis case collections used in the meta-analysis were underpowered to screen for these effects at a modest level of significance ($P < 10^{-4}$). For

example, assuming the observed odds ratios and allele frequencies, simulations show that the rs4810485 SNP in the CD40 gene only had a 53% chance of meeting the $P < 10^{-4}$ significance criteria that we used to initially select SNPs. The other SNPs that replicate would have had only a 19–36% chance of being selected for further replication. Together, this suggests that other common alleles of modest effect size should be identified with additional GWA studies and deeper replication in large autoantibody-positive rheumatoid arthritis sample collections.

Methods

Subject groups

Subject groups are described in detail in Table 1 and in the Supplementary Note. Subjects were subdivided into three separate sets: (i) meta-analysis set, (ii) stage 1 replication set, and (iii) stage 2 replication set. Each collection consisted only of individuals that were self-described white and of European descent, and all cases either met 1987 ACR diagnostic criteria or were diagnosed by board-certified rheumatologists. Informed consent was obtained from each individual, and the institutional review board at each collecting site approved the study.

We used three subject groups to conduct the rheumatoid arthritis GWA meta-analysis. The groups used in the meta-analysis have been described elsewhere, and include those from the WTCCC, NARAC and EIRA. All of the cases in EIRA and NARAC and >80% of cases in WTCCC are CCP positive. For the WTCCC collection, we used an expanded collection of controls drawing from five non-autoimmune diseases that were genotyped as part of the larger study.

We used eight subject groups in our replication set. These collections consisted entirely of cases that were autoantibody positive (CCP or RF). For most of the collections, control samples were collected along with case samples as part of the same study. For some of the collections, control samples were unavailable; we matched these case collections to publicly available shared controls that had been genotyped on compatible platforms. The stage 1 replication set consisted of three subject groups: (i) CCP- or RF-positive cases identified by chart review from the Nurses Health Study (NHS) and matched controls based on age, gender, menopausal status, and hormone use; (ii) CCPpositive cases from the Brigham Rheumatoid Arthritis Sequential Study (BRASS) and controls from the National Institutes of Mental Health (NIMH); and (iii) CCP-positive cases drawn from North American clinics and controls from the New York Cancer Project (together this collection is called NARAC-II). The stage 2 replication set consisted of five collections: (i) CCPpositive cases drawn from North American clinics (NARAC-III) (P.K.G., unpublished data) and publicly available controls taken from a Parkinson's study and study 66 and 67 of the Illumina Genotype Control Database; (ii) North American RF-positive cases and controls matched on gender, age and grandparental country of origin from the Genomics Collaborative Initiative;(iii) CCP- or RF-positive Dutch cases and controls from Leiden University Medical Center (LUMC); (iv) CCP-positive cases from Sweden and epidemiologically matched controls (EIRA-II); and (v) CCP-positive Dutch cases and controls collected from the greater Amsterdam region (GENRA).

Genotyping

Detailed description of genotyping is provided in the Supplementary Note. All GWA meta-analysis genotyping was previously described. We directly genotyped 38 SNPs in stage 1 replication samples with the Sequenom iPlex platform at the Broad Institute (for NHS case-control samples and BRASS case samples) and National Institutes of Health (for NARAC-II and NYCP samples). We obtained NIMH genotypes from previously generated GWA data on the Affymetrix 500K platform through a formal application process. We genotyped stage 2 replication samples with (i) the Illumina 317K array at the Feinstein Institute (for the NARAC-III samples; unpublished data); (ii) using the kinetic PCR platform at Celera Diagnostics (for the GCI and LUMC samples); and (iii) with the Sequenom iPlex platform at the Broad Institute (for the EIRAII and AMC/UVA samples). We obtained publicly available genotype data for shared controls for NARAC-III cases after an official application to a Parkinson's Disease consortium and Illumina Genotype Control Database. All stage 2 SNPs were directly genotyped in the GCI, LUMC, EIRA-II and GENRA samples, and individually imputed in NARAC-III case-control samples to determine genotype probability, as in our GWAS meta-analysis (see below).

In stages 1 and 2 we required that each SNP pass the following criteria for each collection separately: (i) genotype missing rate <10%, (ii) minor allele frequency >1%, and (iii) Hardy-Weinberg equilibrium with $P > 10^{-3}$. We also excluded individuals with data missing for >10% of SNPs. Of the 38 SNPs advanced into stage 1, 6 SNPs failed genotyping in Sequenom iPlex at either the Broad or NIH, and 2 failed in the NIMH dataset. The remaining SNPs had <4% missing data for each collection. All 17 SNPs passed stage 2 replication in NARAC-III, 2 failed in GCI and LUMC, and 4 failed in EIRA-II and GENRA.

Imputation and GWA meta-analysis

We conducted a GWA meta-analysis on a set of SNPs genotyped in the WTCCC study (Supplementary Note). We selectedWTCCC SNPs on the basis of strict quality control criteria: (i) genotype missing rate <1% in cases and controls separately, (ii) minor allele frequency >1% in cases and controls separately, (iii) Hardy-Weinberg equilibrium with $P > 10^{-4}$ by a 2 degree-of-freedom test in cases and controls separately and (iv) availability of Phase II HapMap data. This resulted in a total of 336,721 SNPs. We imputed these SNPs in the EIRA and NARAC collections with IMPUTE. We used EIRA and NARAC data that had been filtered and imputed genotypes separately. We conducted separate runs for each chromosome using default parameters. As IMPUTE provides probabilistic confidence scores that track with prediction accuracy, we elected to use probabilistic dosages in our statistical analysis rather than hard genotype calls. This approach accounts for some uncertainty in imputation, and avoids bias.

To address case-control stratification we used identity-by-state to cluster EIRA cases and controls on the basis of Illumina 317K SNP data into 165 substrata, and then to cluster NARAC cases and controls similarly on the basis of Illumina 550K SNP data into 396 substrata. This strategy was identical to that used to effectively control stratification previously in this dataset. As previous investigations revealed minimal case-control stratification in the WTCCC data, we placed all cases and controls from the WTCCC into a single stratum.We calculated association statistics using genotype counts available online for the WTCCC (see URLs section below) and probabilistic allele dosages for EIRA and NARAC. We calculated a CMH 1 degree of freedom test on the basis of allele frequency across 562 strata, and then after correcting χ^2 scores by genomic control inflation (λ_{GC}), we assigned P values.

Population stratification in replication samples.

For each replication collection we corrected for possible case-control stratification by either (i) using only epidemiologically matched samples when cases and controls were drawn from the same population, or (ii) matching at least one control for each case on the basis of ancestry informative markers (see Supplementary Note for details). As the cases in the NHS, GCI, LUMC, EIRA-II and GENRA collections were well matched to controls, we did not pursue further strategies to correct for case-control stratification. For the BRASS, NARAC-II and NARAC-III collections, we matched cases and controls with ancestry informative markers and placed them into a single stratum. For the BRASS cases and NIMH controls, GWA data on Affymetrix 6.0 (unpublished data) and Affymetrix 500K platforms were available, respectively. A total of 57,417 SNPs overlapped both datasets that had 0% missing data across all individuals; we used these as SNPs to derive ancestry information. For NARAC-II cases and NYCP controls, cases and controls were matched using genotype data on 760 ancestry informative markers. Finally for the NARAC-III cases (unpublished data) and shared controls, we used available Illumina 317K GWA data for 269,771 SNPs passing stringent quality control criteria. For each case-control collection, we used these SNPs to define the top ten principal components and to remove genetically distinct outliers (σ threshold = 6 with five iterations) with the software program EIGENSTRAT. We eliminated vectors that correlated with known structural variants on chromosomes 8 and 17, demonstrated minimal variation, or did not stratify cases and controls. After mapping cases and controls in the space of eigenvectors, we matched cases to controls that were nearest in Euclidean distance. A total of 814 of the available 1,498 NIMH controls were included (matching along the top two principal components), a total of 637 of the available 1,153 NARAC-II controls were included (matching along the top principal component), and a total of 1,303 of the available 2,189 NARAC-III shared controls were included (matching along the top two principal components).

Stage 1 and 2 replication analysis

We selected SNPs for replication by (i) identifying all SNPs that achieve $P < 10^{-4}$ significance in meta-analysis, (ii) grouping SNPs with $r^2 < 0.1$ into regions, and (iii) forwarding the SNP showing strongest association from each region for replication (Supplementary Note). We excluded SNPs from regions that had already demonstrated association in other studies. We forwarded 38 SNPs for stage 1 replication. The most significant SNPs from a preliminary statistical analysis conducted without correcting for possible case-control stratification were forwarded for stage 2 replication. For only those SNPs that replicated with $P < 0.05$, we genotyped EIRA samples and replaced imputed genotypes with genotype data for our final analysis (Supplementary Table 5 online). For each SNP we conducted three statistical tests. First, we conducted a onesided CMH statistical test across eight strata to assess whether rheumatoid arthritis association was reproducible in the replication collections in the same direction as the GWAS meta-analysis used to select the SNPs of interest. Second, we calculated a 570 strata joint analysis across all meta-analysis strata and substrata and replication strata; the eight replication collections were each placed into their own strata and the GWAS samples were partitioned into 562 strata, as described above. We considered $P < 5 \times 10^{-8}$ as a reproducible level of significance. Third, we calculated a Breslow-Day test of heterogeneity of odds ratios. We performed all analyses in MATLAB.

References

(1) The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661–678 (2007).

(2) Plenge, R.M. et al. TRAF1–C5 as a risk locus for rheumatoid arthritis–a genomewide study. N. Engl. J. Med. 357, 1199–1209 (2007).

(3) Kurreeman, F.A. et al. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS Med. 4, e278 (2007).

(4) Plenge, R.M. et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat. Genet. 39, 1477–1482 (2007).

(5) Thomson, W. et al. Rheumatoid arthritis association at 6q23. Nat. Genet. 39, 1431–1433 (2007).

(6) Stastny, P. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. N. Engl. J. Med. 298, 869–871 (1978).

(7) Begovich, A.B. et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am. J. Hum. Genet. 75, 330–337 (2004).

(8) Remmers, E.F. et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N. Engl. J. Med. 357, 977–986 (2007).

(9) Zhernakova, A. et al. Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. Am. J. Hum. Genet. 81, 1284–1288 (2007).

(10) Plenge, R.M. et al. Replication of putative candidate-gene associations with rheumatoid arthritis in 44,000 samples from North America and Sweden: association of susceptibility with PTPN22, CTLA4, and PADI4. Am. J. Hum. Genet. 77, 1044–1060 (2005).

(11) Suzuki, A. et al. Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. Nat. Genet. 34, 395–402 (2003).

(12) Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. Nat. Genet. 39, 906–913 (2007).

(13) Jacobson, E.M. et al. A CD40 Kozak sequence polymorphism and susceptibility to antibody-mediated autoimmune conditions: the role of CD40 tissue-specific expression. Genes Immun. 8, 205–214 (2007).

(14) Heward, J.M. et al. A single nucleotide polymorphism in the CD40 gene on chromosome 20q (GD-2) provides no evidence for susceptibility to Graves' disease in UK Caucasians. Clin. Endocrinol. 61, 269–272 (2004).

(15) Houston, F.A. et al. Role of the CD40 locus in Graves' disease. Thyroid 14, 506–509 (2004).

(16) Kawabe, T. et al. The immune responses in CD40-deficient mice: impaired immunoglobulin class switching and germinal center formation. Immunity 1, 167–178 (1994).

(17) Lougaris, V., Badolato, R., Ferrari, S. & Plebani, A. Hyper immunoglobulin M syndrome due to CD40 deficiency: clinical, molecular, and immunological features. Immunol. Rev. 203, 48–66 (2005).

(18) Manzo, A. et al. Systematic microanatomical analysis of CXCL13 and CCL21 in situ production and progressive lymphoid organization in rheumatoid synovitis. Eur. J. Immunol. 35, 1347–1359 (2005).

(19) Ishida, T. et al. CD40 signaling-mediated induction of Bcl-XL, Cdk4, and Cdk6. Implication of their cooperation in selective B cell growth. J. Immunol. 155, 5527–5535 (1995).

(20) Veiga-Fernandes, H. & Rocha, B. High expression of active CDK6 in the cytoplasm of CD8 memory cells favors rapid division. Nat. Immunol. 5, 31–37 (2004).

(21) Carpenter, C.L. Btk-dependent of phosphoinositide synthesis. Biochem. Soc. Trans. 32, 326–329 (2004).

(22) Marsters, S.A. et al. Herpesvirus entry mediator, a member of the tumor necrosis factor receptor (TNFR) family, interacts with members of the TNFR-associated factor family and activates the transcription factors NF-kappaB and AP-1. J. Biol. Chem. 272, 14029–14032 (1997).

(23) Gruber, T. et al. PKCtheta cooperates with atypical PKCzeta and PKCiota in NF-kappaB transactivation of T lymphocytes. Mol. Immunol. 45, 117–126 (2008).

(24) Barton, A. et al. Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. Nat. Genet. advance online publication, doi:10.1038/ng.218 (14 September 2008).

(25) Harnett, M.M. CD40: a growing cytoplasmic tale. Sci. STKE 2004, pe25 (2004).

(26) Xie, P., Hostager, B.S., Munroe, M.E., Moore, C.R. & Bishop, G.A. Cooperation between TNF receptor-associated factors 1 and 2 in CD40 signaling. J. Immunol. 176, 5388–5400 (2006).

(27) Song, H.Y., Rothe, M. & Goeddel, D.V. The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappaB activation. Proc. Natl. Acad. Sci. USA 93, 6721–6725 (1996).

(28) Balasa, B. et al. CD40 ligand-CD40 interactions are necessary for the initiation of insulitis and diabetes in nonobese diabetic mice. J. Immunol. 159, 4620–4627 (1997).

(29) Durie, F.H. et al. Prevention of collagen-induced arthritis with an antibody to gp39, the ligand for CD40. Science 261, 1328–1330 (1993).

(30) MacGregor, A.J. et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum. 43, 30–37 (2000).

Chapter 16

Association of *IL2RA* and *IL2RB* with rheumatoid arthritis - a replication study in a Dutch population

F.A.S. Kurreeman
N.A. Daha
M. Chang
J.J. Catanese
A.B. Begovich
T.W.J.Huizinga
R.E.M. Toes

Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands.
Celera, Alameda, CA, USA.

Manuscript in press at Ann. Rheum Dis.

Rheumatoid arthritis is an autoimmune-disease with a prevalence of ~1% world wide. The etiology of RA is largely unknown; however it is thought that both genetic as well as environmental factors play a role in the pathogenesis of the disease. Genome wide association studies (GWAS) as well as candidate gene approaches have led to the association of a number of genetic susceptibility loci (1-7). The Wellcome Trust case-control consortium (WTCCC), the first GWAS in RA, identified a number of loci reaching genome-wide significance including the HLA-region and the *PTPN22* gene (5). To identify new genetic risk factors, Thomson *et al*. investigated whether tier 2 single nucleotide polymorphisms (SNPs) ($P=1x10^{-5}$-$1x10^{-7}$) in the WTCCC-GWAS showed association with RA in an independent validation study of 5063 patients and 3849 healthy controls(8). Of the nine loci investigated, significant association with rs6920220 in the *TNFAIP3-OLIG3* region was identified (OR 1.23, 95%CI-1.15-1.33, $P=1.1x10^{-8}$). Association of RA with this region was independently identified by Plenge and coworkers (2) and a recent meta-analysis of GWAS data from three independent case-control populations of European descent confirmed these results(3). To further investigate these tier 2 SNPs, the control group of the validation study was enlarged from 3849 to 11487 individuals by including non-RA disease groups consisting of bipolar disorder, Type 2 Diabetes, hypertension and coronary artery disease(9). In addition to association with rs6920220, other statistically significant SNPs surfaced, located in *MMEL1*, *IL2RA* and *IL2RB*. Furthermore, investigation of forty-nine tier 3 loci ($P=1x10^{-4}$-$1x10^{-5}$) from the original WTCCC-GWAS, identified significant RA-associated SNPs located in *PRKCQ* and *KIF5A* (8). Three of these regions, encoding *MMEL1*, *KIF5A* and *PRKCQ,* were also identified in a large independent study of RA samples of European descent (3), suggesting that these regions harbor true RA susceptibility loci.

In the present study we addressed the contribution of the two IL-2-pathway SNPs, specifically rs743777 located in *IL2RB* and rs2104286 in *IL2RA,* to RA risk in an independent Dutch case-control study. This is of relevance, not only because additional replication would strengthen the putative contribution of IL-2-receptor-positive cells to RA, but also because only a trend towards association for the *IL2RA* SNP, was observed after correction for multiple testing using the Bonferroni-method ($P<0.0056$). 616 Dutch patients and 545 healthy, ethnically and geographically matched controls were genotyped by allele specific kinetic PCR as previously described(10). In this study, significant association with RA for both rs743777 and rs2104286 was observed (OR 1.26, 95%CI 1.06-1.50, P=0.009; OR 0.81, 95%CI 0.67-0.98, P=0.026 respectively) (Table1). Combining the data from our study with the UK data strengthened the evidence for association (rs743444: OR 1.12, 95%CI 1.06-1.18, $P=8.6x10^{-6}$; rs2104286: OR 0.92, 95%CI 0.87-0.97, $P=1.2x10^{-3}$) (Table1). To bypass the phenomenon of the 'winners curse', in which effect sizes are often overestimated in the original study(13), we opted to analyze the data without the original WTCCC data to provide an estimate of the most likely effect size. In conclusion, this study provides additional evidence for the association of *IL2RB* and *IL2RA* with RA by independent replication in a Dutch population, underlining the importance of the IL-2 pathway in RA. Recently the *IL2RA* region was also found to be associated with other autoimmune diseases(11;12), specifically multiple sclerosis and type 1 diabetes, which suggests a possible common functional pathway.

		cases					controls						
	N	11	12	22	MAF	N	11	12	22	MAF	OR(95%CI)	P-value	HW-con
IL2RB *rs743777*													
This study	616	76(*0.12*)	288(*0.47*)	252(*0.41*)	0.36	544	58(*0.11*)	217(*0.40*)	269(*0.49*)	0.31	1.26(1.06-1.50)	0.009	0.37
Validation study UK	4680	532(*0.11*)	2031(*0.43*)	2117(*0.45*)	0.33	11200	1040(*0.09*)	4832(*0.43*)	5328(*0.48*)	0.31	1.11(1.05-1.17)	0.0001	0.50
Combined											**1.12(1.06-1.18)**	**8.6*10-6**	
IL2RA *rs2104286*													
This study	616	36(*0.06*)	226(*0.37*)	354(*0.57*)	0.24	545	54(*0.10*)	200(*0.37*)	291(*0.53*)	0.29	0.81(0.67-0.98)	0.026	0.10
Validation study UK	4660	312(*0.07*)	1740(*0.37*)	2608(*0.56*)	0.24	11260	790(*0.07*)	4464(*0.40*)	6006(*0.53*)	0.28	0.93(0.88-0.98)	0.007	0.59
Combined											**0.92(0.87-0.97)**	**0.0012**	

Table 1. Analysis of rs743777 and rs2104286 with RA in two populations of northern-European descent. Allele frequency data of RA cases versus controls was compared using a fixed effects (Mantel-Haenszel) meta-analysis. No significant heterogeneity or deviation from Hardy-Weinberg (HW) equilibrium was observed among the studies.

Reference List

(1) Kurreeman FA, Padyukov L, Marques RB, Schrodi SJ, Seddighzadeh M, Stoeken-Rijsbergen G et al. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS Med 2007; 4(9):e278.

(2) Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat Genet 2007; 39(12):1477-82.3

(3) Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, Burtt NP et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat Genet 2008; 40(10):1216-23.

(4) Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. N Engl J Med 2007; 357(10):977-86.

(5) The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007; 447(7145):661-78.

(6) Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, Franke B et al. Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. Am J Hum Genet 2007; 81(6):1284-8.

(7) Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC et al. Amissense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatise (PTPN22) is associated with rheumatoid arthritis. Am J Hum Genet 2004; 75(2):330-7.

(8) Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J et al. Rheumatoid arthritis association at 6q23. Nat Genet 2007; 39(12):1431-3.

(9) Barton A, Thomson W, Ke X, Eyre S, Hinks A, Bowes J et al. Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. Nat Genet 2008; 40(10):1156-9.

(10) Chang M, Rowland CM, Garcia VE, Schrodi SJ, Catanese JJ, van der Helm-van Mil AH et al. A large-scale rheumatoid arthritis genetic study identifies association at chromosome 9q33.2. PLoS Genet 2008; 4(6):e1000107.

(11) Vella A, Cooper JD, Lowe CE, Walker N, Nutland S, Widmer B et al. Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms. Am J Hum Genet 2005; 76(5):773-9.

(12) Matesanz F, Fernandez O, Alcina A. Genomewide study of multiple sclerosis. N Engl J Med 2007; 357(21):2200-1.

(13) Kraft P. Curses--winner's and otherwise--in genetic epidemiology. Epidemiology 2008; 19(5):649-51.

Chapter 17

Summary & Discussion

Rheumatoid Arthritis (RA) is a complex disease with unknown etiology. Genetic as well as environmental risk factors are both thought to play a role in disease onset and/or progression[1, 2].

At the immunological level, the role of the adaptive immune system in autoimmune diseases is largely accepted. Based on both mouse models and data from human studies, it has become clear that T cells and B cells, amongst other immune cells, are likely to play a major role in RA. Examples of such studies include the clear role of T cells in the development of RA in the SKG mouse model where a mutation in the ZAP-70 molecule (a key signal transduction molecule in T cells) leads to spontaneous arthritis[3]. Additional evidence for the relevance of T cells in arthritis comes from the very strong association of the *human leukocyte antigen* (HLA) region with RA in human genetic studies[4, 5]. The products of the HLA-DRB1 alleles that associate the strongest with RA share a 5 amino-acid sequence in a peptide-binding pocket called Shared Epitope (SE). The prediction has been that these DRB1 molecules would bind and allow the presentation of RA inducing peptide(s) to T cells[6]. Although these peptides remain elusive, the well-established association of the SE with RA provides evidence of the role of T cells in RA development. However it is becoming increasingly clear that it is not simply a case of SE-positive alleles predisposing to disease. The presence of the SE alleles is highly correlated with the presence of autoantibodies indicating that SE alleles likely do not predispose to the development of the disease as such[7, 8]. Additionally, there may also be HLA-DRB1 protective alleles and also a number of other important associations to polymorphisms in this gene-rich region of the genome6, 9-11.

The role of B cells in RA is highlighted by the presence of autoantibodies in patients. These autoantibodies including rheumatoid factor (RF) and anti-citrullinated protein antibodies (ACPA) have been consistently described12, 13. The emergence of ACPA years before onset of clinical symptoms indicates that it possibly lies in a causal pathway for the development of the disease although no conclusive evidence could be obtained from mouse models14. In 2006, Kuhn and colleagues showed that ACPA can exacerbate disease progression in a collagen-induced mouse model15. In 2009, a recent report suggests that these autoantibodies specific to citrullinated proteins do play a role in both the onset and progression of autoimmunity in the same mouse model16. However, it still takes years before clinical symptoms are apparent in patients who already harbour these autoantibodies, indicating that other genetic as well as non-genetic factors likely play a role in the onset of the disease.

Despite the rather large genetic component in RA (approximately 60%), it has been very difficult to identify the precise genetic markers responsible for the onset of the disease4, 17. The increased risk of disease in siblings of patients with RA compared with that of the general population (λs) has been estimated to be between 2 and 174. However, recent developments in genetics and genomics, heavily influenced by technological innovations, are overcoming the prior limitations within the field.

Genetic variations are sometimes referred to as “polymorphisms”, meaning that the gene or locus (the gene region) occurs in several “forms” within the population. Most loci that are polymorphic have no direct influence on disease risk or human traits, while those that are associated with a difference in risk of disease or a human trait (however subtle) are termed “disease-associated polymorphisms”. Although the clinical significance and causality of these disease-associated polymorphisms are currently difficult to establish, these variations may account for a genetically susceptible individual predisposed to an aberrant immune profile, who,

combined with other non-genetic factors eventually develops diseases like RA[18]. In line with this concept, human genetic studies offer *prima facie* evidence that a biological pathway is critical in disease pathogenesis. Since there are no drugs yet proven to cure RA or other chronic inflammatory/autoimmune diseases, research into the principle mechanisms underlying the disease process are crucial for disease management, new diagnostics and ultimately cure.

In advance of large-scale techniques, which have now become widely available, we took a hypothesis-driven approach to determine whether candidate genes, namely complement component 5 (C5) and interleukin 10 (IL10), both immune related genes, may be involved either in the development and/or progression of RA. This approach has been highly rewarding as exemplified identification of one of the few widely-confirmed genetic risk factors for RA, namely the *Tumour Necrosis Factor (TNF) receptor associated factor 1* (*TRAF1*) – *Complement component 5* (*TRAF1/C5)* locus on chromosome 9q33 (**part I of this thesis)** and by our functional characterization of the mRNA regulation at the *interleukin 10* (*IL10*) locus (**part II of this thesis**). In **part III**, we have also investigated the role of several candidates in relation to disease as well as specific disease phenotypes.

Genetic Risk factors in RA

Prior to 2007, only two genetic risk factors had been identified. The first and most robust genetic risk factor to date was identified three decades ago and is confined to the *HLA* region on chromosome 6[4, 5]. The second largest risk factor in RA, the *protein tyrosine phosphatase non-receptor type 22* (*PTPN22)* gene, was identified by using a largescale approach in 2004 and has now been widely replicated as an autoimmune locus[19-22].

In collaboration with our Swedish and American colleagues, we successfully identified a novel genetic risk factor in RA in 2007 in the 9q33 region of the genome containing *TRAF1/C5* by taking a candidate gene approach (**chapter 2**)[23]. Interestingly, this region was also concurrently detected in a genome-wide approach by Plenge and colleagues[24]. Although the original genome wide association study (GWAS) performed by the Wellcome Trust Case-Control Consortium (WTCCC) in 2000 patients and 3000 controls did not originally identify this locus, a follow-up post-genome study performed in the UK now also provides evidence of association at this region[25, 26].

Our data revealed that although the case–control allele frequency increase in different sample sets ranged from 4% to 9%, the frequency ranged from 38% to 46% in control groups of European ancestry. Data from HapMap (www.hapmap.org) support these observations with the G allele (minor allele in Caucasians) frequency of rs3761847 (perfect proxy of rs10818488) varying from 48% in Caucasians of European descent, 31% in Gujarati Indians in Texas, 42% in Japanese in Tokyo, 66% in Mexicans in Los Angeles, 59% in Yorubans in Nigeria and 74% in Luhya in Kenya. Given that association studies compare the frequencies in patients versus healthy individuals, unknown biases in control frequencies may lead to spurious associations. To address this issue and to further complement our data, we have also reproduced this association in 1356 individuals from 452 trio families with RA (chapter 3)27. Interestingly, an RA association of the opposite allele has been reported in the Japanese population but no association has been found in a Korean population, reflecting possible ethnic diversity in the associations at this locus28, 29. A more recent fine-mapping of this region now indicates that a

third gene, PHD finger protein 19 (PHF19), may be part of the haplotype block involved in disease susceptibility30. While TRAF1 and C5 being immune-related molecules represent ideal candidates as biological mediators in RA, relatively little is currently known about PHF19 to infer its possible role in the disease process.

Our work has now also shown that the *TRAF1/C5* region is not only relevant for RA but is also relevant in patients with a polyarticular form of juvenile arthritis (*Odds ratio* 1.46, *p*=0.004) (**chapter 4**)[31]. Behrens and colleagues have also independently reported an association of a perfect proxy to JIA further supporting our findings[32]. In recent years, several common variants associated with RA have been convincingly associated with multiple autoimmune diseases including the PTPN22[33] and *signal transducer and activator of transcription 4* (*STAT4*)[34] genes.

In line with these observations, one widely accepted pattern across the genes loci identified is the overlap of genetic risk factors across various autoimmune diseases. To test this hypothesis we investigated the *TRAF1/C5* locus in a well-powered study including four additional autoimmune diseases including 735 Type I Diabetes (TID), 1049 Celiac Disease (CD), 367 Systemic Sclerosis (SSc), 746 Systemic Lupus Erythematosus (SLE) patients and a common set of controls consisting of 3494 healthy individuals. In combination with the previously unpublished data from the SLEGEN consortium, we observe an overall association to SLE in 1577 patients and 4215 healthy individuals (OR 1.22, $p=1.02x10^{-6}$). We found suggestive evidence for T1D and did not observe any evidence of association to CD and SSc (**chapter 5**)[35].

With these studies, we have provided considerable evidence that the *TRAF1/C5* region is not only relevant to RA but that the frequency of the same allele is increased in JIA and SLE. It is therefore highly likely that the *TRAF1/C5* region is a genetic risk factor involved in a shared pathway underlying multiple autoimmune diseases. It now seems that most of the RA susceptibility loci identified so far have an association with other autoimmune diseases[36]. Interestingly, the association with SLE is absent in a well-powered Japanese case-control study and small study in the Columbian population[37, 38]. As with many genetic loci, it is highly likely that ethnic differences exist in the contribution of this locus to SLE. Since, *TRAF1* is suggested to be a negative regulator of TNF-Receptor and/or CD40 signaling and *C5* is a central component of the complement pathway and they are both highly involved in the immune system, they both represent likely candidates for further functional studies. Current efforts are geared towards the functional relevance of the genetic variants in the *TRAF1/C5* region by determining the expression levels and function of each gene in primary lymphocytes from healthy individuals to tease apart the influence of these genetic variants on biological function.

In addition to the identification of the *TRAF1/C5* locus, the year 2007 saw the discovery and replication of two additional risk loci, more than doubling the number of genes known to play a role in RA[39]. The *STAT4* gene region on chromosome 2q was identified following a mapping of genes under a linkage peak and so far represents the only linkage study to have successfully pinpointed the gene of interest in RA in Caucasians[40]. The other locus, identified by the WTCCC as well as Plenge and colleagues using a genome wide association study (GWAS) with 100,000 markers is located on chromosome 6q23 and encompasses a very interesting candidate gene *TNF-α induced protein 3* (*TNFAIP3)*[25, 41].

Unprecedented progress however has been witnessed in 2008 when a meta-analysis of three genome-wide association studies from the US, Sweden and the UK was performed (**chapter**

15)[42]. Comprised of 3,393 patients who were predominantly ACPA positive and 12,460 healthy individuals, our study enabled the confirmation of previously associated RA loci (*HLA*, *PTPN22*, *TRAF1/C5*, *STAT4* and *6q23* containing *TNFAIP3*)[39] and also provided suggestive evidence towards previous associations with *Cytotoxic T Lymphocyte associated 4 (CTLA4)* and *interleukin 2/interleukin 21* (*IL2/21)*[43, 44]. By genotyping these variants in our Dutch cohort, we have now independently replicated three of these loci (*STAT4*, *IL2/21* and *CTLA4*) in the Dutch population further confirming the suggested role of *IL2/21* and *CTLA4* in RA (**chapter 12**)[45]. It is of note to mention that an overall estimate of the effect size of the *CTLA4* locus results in an odds ratio of ~1.10, highlighting that lack of power of most previous studies is the likely culprit in the controversy of this locus in RA over the past few years. Both *IL2/21* and *CTLA4* have now been reported in a large UK cohort[36]. In addition, our data also provides evidence of the association of *IL2/21* in juvenile arthritis (**chapter 13**)[46].

The above-mentioned meta-analysis of three GWAS of European-ancestry populations identified 31 regions of interest outside of the known RA risk loci ($P<10^{-4}$). Following replication in 3,929 autoantibody positive individuals (ACPA and/or RF) and 5,807 matched controls, six additional loci were identified with the most significant finding localized in the CD40 gene region, more recently complemented by independent replication[47]. The five additional signals were found in regions harboring the following genes, *membrane metallo-endopeptidase-like 1-Tumor necrosis factor receptor superfamily member 14-TNF receptor superfamily 14* (*MMEL1-TNFRSF14)*, *cyclin-dependent kinase 6* (*CDK6)*, *Chemokine (C-C motif) ligand 21(CCL21)*, *Protein kinase C theta (PRKCQ)*, and *Kinesin family member 5A- Phosphatidylinositol-5-phosphate 4-kinase type II gamma* **(***KIF5A-PIP4K2C)*. A parallel study investigating suggestive hits from the WTCCC study provides additional evidence for associations with *PRKCQ*, *KIF5A-PIP4K2C* and *MMEL1-TNFRSF14* regions in an independent UK sample set[48]. Despite the fact that *CDK6* and *CCL21* have not as yet been independently replicated by other groups, recent data from our group suggests that *CDK6* predisposes to a more severe disease course in ACPA positive patients (Manuscript in Press,A&R). However, additional replication from independent groups will be required to establish whether the *CDK6* and *CCL21* regions are genuine RA loci. Interestingly a recent UK study provides suggestive evidence for *CCL21* but no association with *CDK6*[47]. Barton and colleagues have also suggested a possible role of *IL2RA* and *IL2RB*[48]. We have now confirmed these two loci in our independent Dutch sample set (**chapter 16**) leading to a three-fold increase in the number of RA risk loci in 2008 (Table 1). It will be highly relevant to determine whether gene-gene or gene-environment exists among these loci.

Table 1. Non-HLA susceptibility genes for RA

Gene locus	Chromosomal region	Year of confirmation	Method of identification
PTPN22	1p13	2004	Largescale missense SNP screen
TRAF1-C5	9q33	2007	Candidate Gene Study, GWAS
STAT4	2q33	2007	Linkage
TNFAIP3-OLIG3	6q23	2007	GWAS
IL2-IL21	4q27	2008	Candidate Gene Study
CTLA4	2q33	2008	Candidate Gene Study
CD40	20q13	2008	GWAS
MMEL1-TNFRSF14	1p36	2008	GWAS
KIF5A	12q13	2008	GWAS
PRKCQ	10p15	2008	GWAS
CCL21	9p13	-	GWAS
CDK6	7q21	-	GWAS
IL2RA	10p15	2008	GWAS
IL2RB	22q13	2008	GWAS

Genetic risk factors in phenotypes of RA

RA, which remains a clinical diagnosis based on classification criteria, is most likely not a single entity. This point has been clearly illustrated by dichotomy of ACPA-positive versus ACPA-negative RA with the disease course being more severe in ACPA-positive individuals as compared to ACPA-negative individuals[49]. From a genetic perspective, loci identified seem to also differ between these two disease categories. For example, HLA SE alleles seem to predispose to the development of these autoantibodies themselves while the *PTPN22* locus predominantly predisposes to the development of ACPA-positive disease[7, 50]. However, most studies have so far been underpowered to detect a conclusive distinctive pattern between the two subgroups. The HLA-DR3 alleles have been consistently associated with ACPA-negative RA[51, 52]. Similar associations have now been suggested for the *interferon regulatory factor 5* (*IRF5*) locus, a gene region identified as a risk factor for SLE. In our studies we have shown that *IRF5* has a stronger association with ACPA-negative disease (**chapter 14**), a finding which is also supported by a recent report by Eguez-Gonzalez and colleagues[53, 54]. A recent report suggests a considerable genetic component in the development of ACPA-negative RA as well[55]. It is therefore anticipated that with increasing collaborations between groups this particular question can be addressed with relative ease in the near future.

From these studies it is clear that multiple loci of modest effect are involved in the onset of RA. However, despite our understanding that the subclassification of RA patients based on their autoantibody profile results in a more homogeneous subgroup, RA remains a syndrome comprising several phenotypes. For example, following diagnosis of early RA, approximately 10% of patients undergo natural remission while others develop either mild or severe joint damage over time[56].

Data linking newly identified genetic polymorphisms to disease outcome in RA are only beginning to emerge. Since the early arthritis cohort in Leiden possesses a wealth of information on these particular disease phenotypes, we can ask the clinically relevant question regarding genetic loci that associate with remission as well as severity of joint damage. One of the interesting candidates that has been consistently investigated in radiographic damage in RA is the TNF-α locus. Historically, this cytokine has been of high relevance in RA with TNF-α blockers in use for years already[57]. However at a genetic level, several inconsistent reports exist between this locus and both the susceptibility and severity of RA. Recent GWAS have not identified this region as a susceptibility factor for RA and our recent investigation reveals that it does not predispose to a more severe disease course after stratification for autoantibody status (**chapter 10**)[42, 58]. This is most likely due to the linkage disequilibrium between *TNF-α* and *HLA-DR3* with the latter being highly correlated to ACPA-negative RA generally following a less severe disease course.

More recently we have highlighted the relevance of taking into account these disease phenotypes in genetic studies of RA (**chapter 11**). Two SNPs on chromosome 6q23 near *TNFAIP3* have been associated with susceptibility to rheumatoid arthritis (RA)[25, 41, 59]. While the initial associations were detected in patients with long-standing RA, no association was found in a Swedish early arthritis cohort. Since these sample sets are well controlled for population stratification issues, a likely explanation for this discordance could be the overrepresentation of patients with severe disease in cohorts with long-standing RA. To this end, we analyzed the effect of the 6q23 region (*TNFAIP3*) on the rate of joint destruction in our early arthritis cohort with a mean duration of follow-up of 5 years. Our data are unique as they cover a long period of radiographic follow-up and have been scrutinized for artefacts such as secular trends in treatment intensity. Albeit based on relatively low patient numbers, our data suggest a contribution of the 6q23 region to the rate of joint destruction in ACPA-positive RA, thereby further refining our understanding of the effects exerted by this locus. Replication of our findings in other cohorts is needed. Nonetheless, this is the first study demonstrating such an effect for genetic polymorphisms located outside the HLA-region in ACPA-positive RA patients. More recently, one additional independent protective allele has been described at this locus and requires further confirmation[60]. It would be highly relevant to investigate this additional signal in relation to the joint damage of RA patients.

In **part II** of this thesis, we have also revisited an old cytokine, repeatedly suggested to be associated with radiographic progression in RA. IL10 is a cytokine with key regulatory and anti-inflammatory function involved in the pathogenesis of various diseases[61-64]. Although the large interindividual differences in the production of IL10 have been extensively associated with polymorphisms and haplotypes of the *IL10* gene, surprisingly little evidence existed that this variation was actually dictated by *IL10* haplotypes[65].

Using the technique of allele-specific transcript quantification, the ratio between two alleles (A and G) of the *IL10* gene was characterized in 15 healthy heterozygous individuals. Two groups were identified whereby donors in group 1 exhibited a 1:1 ratio, whereas those in group 2 exhibited a ratio>1 (P<0.0017). We found that donors heterozygous for haplotype *IL10.2* (one of the four ancient *IL10* haplotypes) were only prevalent in the group that showed higher allelic expression ratios. In our study we show that *IL10* alleles are indeed differentially transcribed in cells from heterozygous individuals and that *IL10* haplotypes dictate production of IL10. These findings showed, for the first time, that interindividual differences in IL10 protein levels could be partially explained at the allele-specific transcriptional level (**chapter 7**)[65]. More recently, Sharma and colleagues have reported the post-transcriptional regulation of IL10 via the binding of microRNA hsa-miR-106a to the 3' UTR in several cell lines[66]. It would be interesting to determine whether this microRNA has an allele-specific effect in primary cells and whether this post-transcriptional regulation of IL10 is functionally relevant in diseases like RA.

In a previous study, an increased risk of familial osteoarthritis (OA) at multiple site was detected in subjects with a low innate production of *IL10* as measured by the same *ex vivo* whole blood assay using lipopolysaccharide stimulation as in several other studies[67, 68]. This finding implicating IL10 in cartilage destruction led us to investigate whether genetic variation in *IL10* contributes to the susceptibility of OA (**chapter 6**). Our study has failed to detect those differences. However, due to the highly limited power of the study, we can only preclude a large effect size of this locus in the development of osteoarthritis. To enable more conclusive results with regards to modest effect sizes, more extensive studies with larger sample sizes will have to be conducted.

In RA, the -A2849G (rs6703630), an *IL10* promoter SNP, was shown to be associated with differences in titres of autoantibodies (RF and ACPA). Moreover the rate of joint destruction in RA patients from the early arthritis cohort was twice as high in patients that were –2849G carrier to those who were not (median rate per year 8 versus 4 Sharp van der Heijde units on X-rays of hand and feet)[69]. However, none of the known *IL10* promoter polymorphisms alone could be a causal variant as shown from our functional study. A more likely possibility of causal mutations is either unknown or untyped SNP(s) located on haplotype *IL10*.2. As the length of the haplotype block around *IL10* is highly relevant to the search for the functional polymorphism(s), we characterized the level of linkage disequilibrium in a region of 217 kb, encompassing *IL10* as well as its neighbouring homologues (*IL19*, *IL20* and *IL24*). We successfully excluded the neighboring genes as potential candidates for harboring the functional cis-acting variants (**chapter 8**)[70]. We further fine-mapped the immediate *IL10* region by genotyping 43 SNPs in 57 healthy unrelated individuals of European descent. One haplotype block was identified restricted to 17 kb encompassing the coding region of the *IL10* gene (5kb) as well as 5'and 3' untranslated regions. HapMap data showed similar linkage disequilibrium patterns. Six tagging SNPs, explaining 93% of the variation in the *IL10* region were identified in our study. We have now genotyped these SNPs in our extensive and clinically well-defined RA cohort to determine their relevance to phenotypes of RA. However, our data show no significant differences in the rate of joint damage and remission with respect to the *IL10* genotypes and haplotypes irrespective of their ACPA status (**chapter 9**). These data suggest that *IL10* polymorphisms most likely do not play a major role in phenotypes of R

More recently however, a GWAS in ulcerative colitis and a meta-analysis of T1D GWAS has identified a signal in the *IL10* gene (rs3024505) located in the 3' UTR[71, 72]. Unfortunately, this polymorphism was not tagged in our fine-mapping approach. While it is clear from the lack of association of this region in the recent GWAS meta-analysis that *IL10* most likely does not predispose to the development of RA as such, it remains important to perform a more extensive sequencing and genotyping effort of this region to definitively exclude any potential role of this locus in RA phenotype.

Translating genetic findings into functionally relevant pathways

One common theme across most loci identified is the fact that most loci either contain multiple genes or they are located in regions devoid of any known genes. However, most loci do harbor at least one candidate gene involved in immune function and thereby do provide important information in potential biological processes. For example, in RA the role of T and B cell functions in general have been highlighted as well as a significant indication of the involvement of the NF-kB signalling pathway. The genes involved in those pathways are discussed below. We believe that many of the proteins encoded by these genes either represent drug targets that have been or will be of high therapeutic potential in the near future. Given the overlap of genetic risk factors across various autoimmune genes, these drug targets may prove useful for more than one disease entity[18, 73]. I have attempted to outline the potential biological relevance of the genetic regions identified by the likely candidates in each region as we know it now but fully aware that this landscape is rapidly evolving.

T cell function

The strongest genetic risk factor for RA is the **SE** alleles from the HLA locus. Activation of the T cell occurs via the engagement of the T cell receptor (TCR) and CD28 on the T cell to the HLA-complex presenting the specific peptides as well as the B7 molecules. In these initial steps, **PRCKQ**, also known as protein kinase C theta (PKCθ) and involved in the phosphorylation of a wide variety of protein targets may play an essential role in relocating to the immunological synapse between the T cell and the antigen-presenting cell (APC) during antigen specific interactions[74, 75]. **PTPN22**, the second strongest genetic risk factor in RA, is involved in regulating the threshold of T cell activation through increased phosphatase activity of downstream proteins that associate with the TCR[21]. **CTLA4** is known to compete with CD28 to limit the extent of T cell activation[76, 77]. Once a T cell is activated, **IL2RA** is induced, enabling the trimerisation of the IL2RA, **IL2RB** and IL2RG to form the high affinity IL2 receptor[78]. **IL2** binds this receptor and through further downstream signaling enhances the proliferation and survival of T cells[79]. Interestingly, **STAT4** and **IL21** are involved in T helper 1 and T helper 17 cells, which are T cell lineages involved in inflammatory reactions[80-82]. IL21 as well as **CDK6** also seem to play a role in the proliferation of activated T cells[83, 84]. More importantly, in the last couple of years, a key role for IL21 in the differentiation of B cells into antibody secreting cells has been reported as a result of T-B cell interaction[85-87]. Interestingly, **CCL21** is also a molecule involved in the homing of lymphocytes to secondary lymphoid organs[88]. Expression of this chemokine is associated with ectopic lymphoid structures and has been implicated in the organization of lymphoid tissue affected by rheumatoid arthritis[89].

Taken together, it seems that these genetic risk factors now identified in RA play a crucial role in the regulation of T cell receptor activation and proliferation, migration as well as differentiation of T cells. It will therefore be interesting to determine how these genetic variations alter their respective gene functions. In the case of PTPN22, strong indications already exist that the susceptibility allele results in a higher threshold of T cell activation which may reduce the negative selection of autoreactive T cells in the thymus[22]. In line with these findings and the fact that T cells may play a more regulatory role by providing help for the activation of B cells, it is not surprising that therapies targeted towards T cells have so far not been proven to be highly effective in RA patients[90]. However, gaining insight into the biological implications will inevitably lead to a better understanding of disease processes.

B cell function

B cells are an integral part of our immune system which is geared towards responses against invading antigens. The genes identified by human genetic studies have now also highlighted a role of certain molecules that play an important role in the activation of B cells through the regulation of both the B cell receptor (BCR) function as well as the second signal required to activate B cells i.e the CD40 signalling cascade. In addition to its role in the increased threshold for T cell activation, the **PTPN22** susceptibility allele has recently been implicated in the reduction of BCR signaling[91, 92]. Associations near **TRAF1**, **TNFAIP3** (A20) and **CD40** itself already suggest the possibility that the CD40 signaling pathway mediates rheumatoid pathogenesis through NF-kB activation[93]. In particular, TRAF1 binds the CD40 receptor and cooperates with TRAF2 to activate NF-kB[94-97]. TRAF1 also binds TNFAIP3, which is a negative regulator of NF-kB signaling[98, 99]. Furthermore, CD40 stimulation results in B cell proliferation through regulation of **CDK6** expression[100]. The CD40 signaling pathway has been investigated in drug development, and mouse models have demonstrated that its disruption could prevent development of immune-mediated arthritis[101]. In addition, the production of RF autoantibodies have been shown to be dependent on CD40 signalling[102]. More recently, the NF-kB family of transcription factors (also known as REL) has been implicated in RA by an extensive GWAS performed by Gregersen and colleagues[103]. The various members of *REL* are *c-Rel*, *p65* (*Rel-A*), *Rel-B*, *p50* (*NFkB-1*) and *p52* (*NFkB-2*) and they have been shown to play important roles in immune responses and autoimmunity[104]. Notably, c-Rel has been reported to physically interact with nuclear CD40 and results in the transcriptional regulation of several target genes[105]. These data further uphold the crucial role of the NFkB pathway as well as the CD40 pathway in RA.

Another potential role of the TNF receptor superfamily has been highlighted by the identification of **TNFRSF14**, the cytoplasmic region of which was found to bind to several TRAF family members, which may mediate the signal transduction pathways that activate the immune response[106]. Although the specific functions of phosphatylinositol 4 phosphate 5 kinase type II γ, **PIP4K2C**, remain largely uncharacterized, it is known that it forms part of a family of kinases which are also thought to play a role in the phosphorylation of proteins downstream of the TNF-receptor signaling cascade[107, 108]. The protein is found in sera from RA patients as well as healthy individuals[109].

In addition to its function as a homing signal for T cells to the lymph nodes, it has been suggested that **CCL21** may also play a role in B cells as it is the high-affinity ligand for CCR7, a

molecule also expressed by B cells, although its precise role in these cells remain to be investigated thoroughly[88].

Taken together, the genes identified seem to play important roles in B cell function, either by regulating the first signal which is the BCR activation or by affecting the second signal which is the signaling cascade resulting from TNF-receptor signaling, in particular via the CD40 receptor.

Altogether, several processes have been highlighted involving antigen recognition, lymphocyte migration, receptor activation, cell fate of T and B cells and NFkB signaling. While the identification of these genetic risk factors have provided considerable insight into the crucial role of the adaptive immune system in RA, we realize that much work lies ahead in the identification of the actual causal variants through sequencing, fine-mapping and functional experiments.

The next step

Common Variants

While pinpointing these gene regions has been a monumental advancement on our previous knowledge of genetic risk factors for RA, the exact common causal mutation(s) is not known for most of these genetic regions. It is therefore paramount to now take steps to characterize the regions of association in more detail, especially since only one out of three common polymorphisms in the human genome have so far been identified[110, 111]. Additionally, recent studies have highlighted the existence of multiple signals at one locus indicating the complexity of associations with disease. One particular example in RA are the multiple alleles at the *TNFAIP3* locus which shows complex haplotypic associations[41, 59, 60] . To resolve these issues across loci identified in RA, we will need to sequence all the regions identified to capture all unknown variations, followed by an in-depth evaluation of association signals conferred by these polymorphisms. Undoubtedly, applying this strategy across multiple autoimmune diseases as well diverse ethnic populations will be invaluable tools in identifying causal alleles and understanding pathways.

Rare variants

The entirety of RA susceptibility variants identified so far does not explain the entire genetic burden of RA, indicating that more remain to be discovered. While current studies in RA are biased towards common variants, increasing evidence points towards the relevance of rare mutations in other complex diseases. Examples of such rare variants influencing common traits exists in Systemic Lupus Erythematosus (*TREX1*)[112], Inflammatory Bowel Disease (*NOD2* and *IL23R*)[113-115], disorders of cholesterol metabolism (*PCSK9*, *ANGPTL4*, among others)[116-119] and deregulated blood pressure levels (SLC12A3, SLC12A1 and KCNJ1)[120]. Notably, numerous examples of non-RA genes exist containing both common mutations of small effect size together with independent rare functional mutations of larger effect size. The most compelling example includes the rare gain- and loss-of-function missense mutations in *PCSK9* which cause deregulated levels of cholesterol while common *PCSK9* variants reproducibly associate with plasma LDL cholesterol levels[116, 121, 122]. An additional 18 loci containing common variants influencing levels of LDL/HDL cholesterol and/or triglycerides was recently reported[123]. Strikingly, rare variants at 9 of these loci have been shown to cause rare Mendelian forms of

dyslipidemia. It is therefore likely that rare mutations contributing to the RA phenotype exist and contribute significantly towards disease pathogenesis. To enable their discovery, sequencing of exon-coding regions in thousands of patients and controls will be required. Compared with common variants, the technologies to detect these rare variants, termed next-generation sequencing, have only recently become available. Hence, our ability to assay them reliably is still maturing but holds great promise for future investigations of our complex genome[124-126].

Structural variations

Unlike single nucleotide polymorphisms, structural variations have largely been ignored in the genetic mapping of common diseases. Although these have been historically difficult to assay until recently, sequencing of several human genomes have now revealed that they comprise over 20% of genetic variants in the human genome[127]. According to Frazer and colleagues, structural variation refers to all base pairs that differ between individuals and that are not single nucleotide variants. These variations include insertion–deletions (indels), block substitutions, inversions of DNA sequences and copy number differences. Interestingly, recent studies that have looked for associations between rare structural variants and autism and schizophrenia have identified specific deletions involved in both of these diseases[128-130]. Additionally, recent technological developments have also generated the possibility of interrogating the genome for copy number variation in disease[131]. Copy numbers in Fc gamma receptor genes have been implicated in SLE as well as RA[132, 133]. We therefore expect that the coming years will provide insight into the relevance of these structural variations in RA and hopefully explain some of the large proportion of the genetic variance in this disease that remains unexplained.

Functional characterization of both common and rare variants

The next frontier lies in the functional characterization of the myriad of sequence variations that influence disease susceptibility. For common variants as well as rare variants, most focus will likely be on candidate protein coding genes and their putative regulatory regions. Experiments are currently being conducted that simultaneously assay global gene expression and genome-wide variation in a large number of individuals to map genetic factors underlying differences in expression levels[134]. While these data sets may undoubtedly be valuable tools for identifying the causative variants and biological bases for many loci associated with a complex trait, they will primarily serve as a prioritization tool for candidate genes and will aid in the hypothesis-generating exercise prior to functional testing. In my opinion, each gene region will require tailor-made experiments in the relevant cell-types to enable a successful translation of genetic findings to biology. Additionally, since each region often encompasses large regions of linkage disequilibrium, there is every possibility that many of these variations will alter the interactions between regulatory (non-coding) RNAs and their targets, a prospect that should not be excluded from future functional analyses[135].

Shared genetic risk factors across diseases and phenotypes – a powerful tool

While sequencing, fine-mapping and identifying common and rare functional variants in RA are the obvious steps the genetic community needs to take, there is one theme in common diseases that can immensely enrich our understanding of the pathways underlying disease. A recent review by Zhernakova and colleagues has recently highlighted the importance of investigating which genetic risk factors are shared by several immune-mediated diseases. The authors investigated genes involved in ankylosing spondylitis, asthma, Graves' disease, celiac disease, Crohn's disease, multiple sclerosis, psoriasis, RA, SLE, type 1 diabetes, and ulcerative colitis and observed that the most likely function of the shared genes were two common pathways involving T cell differentiation and signaling as well as the innate immune response[73]. In addition, a study performed by Smyth and colleagues has highlighted common as well as distinct genetic risk factors between type 1 diabetes and celiac disease[136]. Of the fifteen previously validated susceptibility alleles in type 1 diabetes, at least two contributed to a risk of celiac disease with five showing highly suggestive associations. Interestingly two shared alleles actually had opposite effects in the two diseases: the minor alleles of *IL18RAP* on chromosome 2q12 and *TAGAP* on chromosome 6q25 conferred protection against type 1 diabetes but susceptibility to celiac disease. The authors suggest that common biologic mechanisms, such as autoimmunity related tissue damage and intolerance to dietary antigens may be etiologic features of both diseases.

The concept of shared autoimmunity is now well-established and represents an opportunity not only to study which mechanisms are relevant to several diseases but also which phenotype of a particular disease may have distinct or common pathways. In RA, most of the genetic loci also associate with at least one other autoimmune disease, implying that once a systematic search for these loci has been performed in other immune-related conditions, we can begin to dissect the various overlapping and common mechanisms involved. More recently, suggestions have been made for a multi-step process involving the study of risk factors for specific common sub-phenotypes in diseases, the study of isolated subphenotypes as well as a meta-analysis of these to maximize statistical power in fine-mapping efforts (Huizinga and Grondal, in press at A&R). Therefore combining datasets with well defined phenotypes of each disease may lead to new understanding of disease pathogenesis.

In summary, the advent of technological advances enabling the study of hundreds of thousands of markers across the genome in thousands of cases and controls has significantly changed the genetic landscape of diseases like RA. Now with the sequencing of a thousand genomes (www.1000genomes.org) the genetic community can expect advances on the evaluation of previously unaddressed rare and structural variants in health and disease potentially leading to an even greater understanding of the role of variation in common diseases. In addition, great progress is expected in the coming years in translating these genetic findings into functionally relevant targets for therapeutic interventions.

Reference List

(1) Firestein,G.S. Evolving concepts of rheumatoid arthritis. Nature 423, 356-361 (2003).

(2) Oliver,J.E. & Silman,A.J. Risk factors for the development of rheumatoid arthritis. Scand. J. Rheumatol. 35, 169-174 (2006).

(3) Sakaguchi,N. et al. Altered thymic T-cell selection due to a mutation of the ZAP-70 gene causes autoimmune arthritis in mice. Nature 426, 454-460 (2003).

(4) Seldin,M.F., Amos,C.I., Ward,R., & Gregersen,P.K. The genetics revolution and the assault on rheumatoid arthritis. Arthritis Rheum. 42, 1071-1079 (1999).

(5) Deighton,C.M., Walker,D.J., Griffiths,I.D., & Roberts,D.F. The contribution of HLA to rheumatoid arthritis. Clin. Genet. 36, 178-182 (1989).

(6) de Vries,R.R., Huizinga,T.W., & Toes,R.E. HLA and RA revisited: citrullinated food for the SE hypothesis, the DR6 effect, and NIMA. Hum. Immunol. 67, 454-459 (2006).

(7) van der Helm-van Mil AH et al. The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. Arthritis Rheum. 54, 1117-1121 (2006).

(8) Reveille,J.D. The genetic basis of autoantibody production. Autoimmun. Rev. 5, 389-398 (2006).

(9) Ding,B. et al. Different patterns of associations with anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis in the extended major histocompatibility complex region. Arthritis Rheum. 60, 30-38 (2009).

(10) Coenen,M.J. & Gregersen,P.K. Rheumatoid arthritis: a view of the current genetic landscape. Genes Immun. 10, 101-111 (2009).

(11) Lee,H.S. et al. Several regions in the major histocompatibility complex confer risk for anti-CCP-antibody positive rheumatoid arthritis, independent of the DRB1 locus. Mol. Med. 14, 293-300 (2008).

(12) Shmerling,R.H. & Delbanco,T.L. The rheumatoid factor: an analysis of clinical utility. Am. J. Med. 91, 528-534 (1991).

(13) Schellekens,G.A. et al. The diagnostic properties of rheumatoid arthritis antibodies recognizing a cyclic citrullinated peptide. Arthritis Rheum. 43, 155-163 (2000).

(14) van,G.F., Ioan-Facsinay,A., Huizinga,T.W., & Toes,R.E. The devil in the details: the emerging role of anticitrulline autoimmunity in rheumatoid arthritis. J. Immunol. 175, 5575-5580 (2005).

(15) Kuhn,K.A. et al. Antibodies against citrullinated proteins enhance tissue injury in experimental autoimmune arthritis. J. Clin. Invest 116, 961-973 (2006).

(16) Uysal,H. et al. Structure and pathogenicity of antibodies specific for citrullinated collagen type II in experimental arthritis. J. Exp. Med. 206, 449-462 (2009).

(17) MacGregor,A.J. et al. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. Arthritis Rheum. 43, 30-37 (2000).

(18) Plenge,R.M. Shared genetic risk factors for type 1 diabetes and celiac disease. N. Engl. J. Med. 359, 2837-2838 (2008).

(19) Begovich,A.B. et al. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am. J. Hum. Genet. 75, 330-337 (2004).

(20) Carlton,V.E. et al. PTPN22 genetic variation: evidence for multiple variants associated with rheumatoid arthritis. Am. J. Hum. Genet. 77, 567-581 (2005).

(21) Gregersen,P.K. Gaining insight into PTPN22 and autoimmunity. Nat. Genet. 37, 1300-1302 (2005).

(22) Gregersen,P.K., Lee,H.S., Batliwalla,F., & Begovich,A.B. PTPN22: setting thresholds for autoimmunity. Semin. Immunol. 18, 214-223 (2006).

(23) Kurreeman,F.A. et al. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. PLoS. Med. 4, e278 (2007).

(24) Plenge,R.M. et al. TRAF1-C5 as a risk locus for rheumatoid arthritis - A genomewide study. New England Journal of Medicine 357, 1199-1209 (2007).

(25) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661-678 (2007).

(26) Barton,A. et al. Re-evaluation of putative rheumatoid arthritis susceptibility genes in the post-genome wide association study era and hypothesis of a key pathway underlying susceptibility. Hum. Mol. Genet. 17, 2274-2279 (2008).

(27) Kurreeman,F.A. et al. Replication of the tumor necrosis factor receptor-associated factor 1/complement component 5 region as a susceptibility locus for rheumatoid arthritis in a European family-based study. Arthritis Rheum. 58, 2670-2674 (2008).

(28) Nishimoto,K. et al. Association study of TRAF1-C5 polymorphisms with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in Japanese. Ann. Rheum. Dis.(2009).

(29) Lee,H.S. et al. Genetic risk factors for rheumatoid arthritis differ in Caucasian and Korean populations. Arthritis Rheum. 60, 364-371 (2009).

(30) Chang,M. et al. A large-scale rheumatoid arthritis genetic study identifies association at chromosome 9q33.2. PLoS. Genet. 4, e1000107 (2008).

(31) Albers,H.M. et al. The TRAF1/C5 region is a risk factor for polyarthritis in Juvenile Idiopathic Arthritis. Ann. Rheum. Dis.(2008).

(32) Behrens,E.M. et al. Association of the TRAF1-C5 locus on chromosome 9 with juvenile idiopathic arthritis. Arthritis Rheum. 58, 2206-2207 (2008).

(33) Gregersen,P.K. & Behrens,T.W. Genetics of autoimmune diseases--disorders of immune homeostasis. Nat. Rev. Genet. 7, 917-928 (2006).

(34) Remmers,E.F. et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. New England Journal of Medicine 357, 977-986 (2007).

(35) Kurreeman,F.A. et al. The TRAF1-C5 region on chromosome 9q33 is associated with multiple autoimmune diseases. Ann. Rheum. Dis.(2009).

(36) Barton,A. et al. Identification of AF4/FMR2 family, member 3 (AFF3) as a novel rheumatoid arthritis susceptibility locus and confirmation of two further pan-autoimmune susceptibility genes. Hum. Mol. Genet. 18, 2518-2522 (2009).

(37) Nishimoto,K. et al. Association study of TRAF1-C5 polymorphisms with susceptibility to rheumatoid arthritis and systemic lupus erythematosus in Japanese. Ann. Rheum. Dis.(2009).

(38) Palomino-Morales,R.J. et al. STAT4 but not TRAF1/C5 variants influence the risk of developing rheumatoid arthritis and systemic lupus erythematosus in Colombians. Genes Immun. 9, 379-382 (2008).

(39) Bowes,J. & Barton,A. Recent advances in the genetics of RA susceptibility. Rheumatology. (Oxford) 47, 399-402 (2008).

(40) Remmers,E.F. et al. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. New England Journal of Medicine 357, 977-986 (2007).

(41) Plenge,R.M. et al. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nature Genetics 39, 1477-1482 (2007).

(42) Raychaudhuri,S. et al. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. Nat. Genet. 40, 1216-1223 (2008).

(43) Zhernakova,A. et al. Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. Am. J. Hum. Genet. 81, 1284-1288 (2007).

(44) Plenge,R.M. et al. Replication of putative candidate-gene associations with rheumatoid arthritis in > 4,000 samples from North America and Sweden: Association of susceptibility with PTPN22, CTLA4, and PADI4. American Journal of Human Genetics 77, 1044-1060 (2005).

(45) Daha,N.A. et al. Confirmation of STAT4, IL2/IL21, and CTLA4 polymorphisms in rheumatoid arthritis. Arthritis Rheum. 60, 1255-1260 (2009).

(46) Albers,H.M. et al. Association of the autoimmunity locus 4q27 with juvenile idiopathic arthritis. Arthritis Rheum. 60, 901-904 (2009).

(47) Orozco,G. et al. Association of CD40 with rheumatoid arthritis confirmed in a large UK case-control study. Ann. Rheum. Dis.(2009).

(48) Barton,A. et al. Rheumatoid arthritis susceptibility loci at chromosomes 10p15, 12q13 and 22q13. Nat. Genet. 40, 1156-1159 (2008).

(49) van der Helm-van Mil AH, Verpoort,K.N., Breedveld,F.C., Toes,R.E., & Huizinga,T.W. Antibodies to citrullinated proteins and differences in clinical progression of rheumatoid arthritis. Arthritis Res. Ther. 7, R949-R958 (2005).

(50) Gregersen,P.K. Pathways to gene identification in rheumatoid arthritis: PTPN22 and beyond. Immunol. Rev. 204, 74-86 (2005).

(51) Verpoort,K.N. et al. Association of HLA-DR3 with anti-cyclic citrullinated peptide antibody-negative rheumatoid arthritis. Arthritis Rheum. 52, 3058-3062 (2005).

(52) Irigoyen,P. et al. Regulation of anti-cyclic citrullinated peptide antibodies in rheumatoid arthritis: contrasting effects of HLA-DR3 and the shared epitope alleles. Arthritis Rheum. 52, 3813-3818 (2005).

(53) Sigurdsson,S. et al. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. Arthritis Rheum. 56, 2202-2210 (2007).

(54) eguez-Gonzalez,R. et al. Association of interferon regulatory factor 5 haplotypes, similar to that found in systemic lupus erythematosus, in a large subgroup of patients with rheumatoid arthritis. Arthritis Rheum. 58, 1264-1274 (2008).

(55) van der,W.D. et al. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. Arthritis Rheum. 60, 916-923 (2009).

(56) van der Helm-van Mil AH, Breedveld,F.C., & Huizinga,T.W. Aspects of early arthritis. Definition of disease states in early arthritis: remission versus minimal disease activity. Arthritis Res. Ther. 8, 216 (2006).

(57) Brennan,F.M. & McInnes,I.B. Evidence that cytokines play a role in rheumatoid arthritis. J. Clin. Invest 118, 3537-3545 (2008).

(58) van der Helm-van Mil AH, Kurreeman,F.A., Toes,R.E., & Huizinga,T.W. Association of tumor necrosis factor alpha polymorphism and radiographic progression in rheumatoid arthritis: comment on the article by Khanna et al. Arthritis Rheum. 56, 1032-1033 (2007).

(59) Thomson,W. et al. Rheumatoid arthritis association at 6q23. Nat. Genet. 39, 1431-1433 (2007).

(60) Orozco,G. et al. Combined effects of three independent SNPs greatly increase the risk estimate for RA at 6q23. Hum. Mol. Genet.(2009).

(61) Groux,H. & Cottrez,F. The complex role of interleukin-10 in autoimmunity. J. Autoimmun. 20, 281-285 (2003).

(62) Westendorp,R.G., van Dunne,F.M., Kirkwood,T.B., Helmerhorst,F.M., & Huizinga,T.W. Optimizing human fertility and survival. Nat. Med. 7, 873 (2001).

(63) Westendorp,R.G. et al. Genetic influence on cytokine production and fatal meningococcal disease. Lancet 349, 170-173 (1997).

(64) Mosser,D.M. & Zhang,X. Interleukin-10: new perspectives on an old cytokine. Immunol. Rev. 226, 205-218 (2008).

(65) Eskdale,J. et al. Interleukin 10 secretion in relation to human IL-10 locus haplotypes. Proc. Natl. Acad. Sci. U. S. A 95, 9465-9470 (1998).

(66) Sharma,A. et al. Posttranscriptional regulation of interleukin-10 expression by hsa-miR-106a. Proc. Natl. Acad. Sci. U. S. A 106, 5761-5766 (2009).

(67) Riyazi,N. et al. Association of the risk of osteoarthritis with high innate production of interleukin-1beta and low innate production of interleukin-10 ex vivo, upon lipopolysaccharide stimulation. Arthritis Rheum. 52, 1443-1450 (2005).

(68) Kurreeman,F.A., Schonkeren,J.J., Heijmans,B.T., Toes,R.E., & Huizinga,T.W. Transcription of the IL10 gene reveals allele-specific regulation at the mRNA level. Hum. Mol. Genet. 13, 1755-1762 (2004).

(69) Lard,L.R. et al. Association of the -2849 interleukin-10 promoter polymorphism with autoantibody production and joint destruction in rheumatoid arthritis. Arthritis Rheum. 48, 1841-1848 (2003).

(70) Monraats,P.S. et al. Interleukin 10: a new risk marker for the development of restenosis after percutaneous coronary intervention. Genes Immun. 8, 44-50 (2007).

(71) Barrett,J.C. et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat. Genet.(2009).

(72) Franke,A. et al. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. Nat. Genet. 40, 1319-1323 (2008).

(73) Zhernakova,A., van Diemen,C.C., & Wijmenga,C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. Nat. Rev. Genet. 10, 43-55 (2009).

(74) Altman,A. & Villalba,M. Protein kinase C-theta (PKCtheta): it's all about location, location, location. Immunol. Rev. 192, 53-63 (2003).

(75) Cronin,S.J. & Penninger,J.M. From T-cell activation signals to signaling control of anti-cancer immunity. Immunol. Rev. 220, 151-168 (2007).

(76) Rudd,C.E. & Schneider,H. Unifying concepts in CD28, ICOS and CTLA4 co-receptor signalling. Nat. Rev. Immunol. 3, 544-556 (2003).

(77) Rudd,C.E. The reverse stop-signal model for CTLA4 function. Nat. Rev. Immunol. 8, 153-160 (2008).

(78) Wang,X., Rickert,M., & Garcia,K.C. Structure of the quaternary complex of interleukin-2 with its alpha, beta, and gammac receptors. Science 310, 1159-1163 (2005).

(79) Hoyer,K.K., Dooms,H., Barron,L., & Abbas,A.K. Interleukin-2 in the development and control of inflammatory disease. Immunol. Rev. 226, 19-28 (2008).

(80) Boniface,K., Blom,B., Liu,Y.J., & de Waal,M.R. From interleukin-23 to T-helper 17 cells: human T-helper cell differentiation revisited. Immunol. Rev. 226, 132-146 (2008).

(81) Palmer,M.T. & Weaver,C.T. Immunology: narcissistic helpers. Nature 448, 416-418 (2007).

(82) Holmdahl,R. IL-21 and autoimmune disease--hypothesis and reality? Eur. J. Immunol. 38, 1800-1802 (2008).

(83) Malumbres,M. et al. Mammalian cells cycle without the D-type cyclin-dependent kinases Cdk4 and Cdk6. Cell 118, 493-504 (2004).

(84) Ostiguy,V., Allard,E.L., Marquis,M., Leignadier,J., & Labrecque,N. IL-21 promotes T lymphocyte survival by activating the phosphatidylinositol-3 kinase signaling cascade. J. Leukoc. Biol. 82, 645-656 (2007).

(85) Ettinger,R. et al. IL-21 induces differentiation of human naive and memory B cells into antibody-secreting plasma cells. J. Immunol. 175, 7867-7879 (2005).

(86) Ettinger,R., Kuchen,S., & Lipsky,P.E. The role of IL-21 in regulating B-cell function in health and disease. Immunol. Rev. 223, 60-86 (2008).

(87) Kuchen,S. et al. Essential role of IL-21 in B cell activation, expansion, and plasma cell generation during CD4+ T cell-B cell collaboration. J. Immunol. 179, 5886-5896 (2007).

(88) Forster,R., valos-Misslitz,A.C., & Rot,A. CCR7 and its ligands: balancing immunity and tolerance. Nat. Rev. Immunol. 8, 362-371 (2008).

(89) Manzo,A. et al. Systematic microanatomical analysis of CXCL13 and CCL21 in situ production and progressive lymphoid organization in rheumatoid synovitis. Eur. J. Immunol. 35, 1347-1359 (2005).

(90) Smolen,J.S. & Steiner,G. Therapeutic strategies for rheumatoid arthritis. Nat. Rev. Drug Discov. 2, 473-488 (2003).

(91) Arechiga,A.F. et al. Cutting edge: The PTPN22 allelic variant associated with autoimmunity impairs B cell signaling. J. Immunol. 182, 3343-3347 (2009).

(92) Rieck,M. et al. Genetic variation in PTPN22 corresponds to altered function of T and B lymphocytes. J. Immunol. 179, 4704-4710 (2007).

(93) Tsukamoto,N., Kobayashi,N., Azuma,S., Yamamoto,T., & Inoue,J. Two differently regulated nuclear factor kappaB activation pathways triggered by the cytoplasmic tail of CD40. Proc. Natl. Acad. Sci. U. S. A 96, 1234-1239 (1999).

(94) Rothe,M., Wong,S.C., Henzel,W.J., & Goeddel,D.V. A novel family of putative signal transducers associated with the cytoplasmic domain of the 75 kDa tumor necrosis factor receptor. Cell 78, 681-692 (1994).

(95) Bishop,G.A., Moore,C.R., Xie,P., Stunz,L.L., & Kraus,Z.J. TRAF proteins in CD40 signaling. Adv. Exp. Med. Biol. 597, 131-151 (2007).

(96) Zapata,J.M. & Reed,J.C. TRAF1: lord without a RING. Sci. STKE. 2002, E27 (2002).

(97) Hu,H.M., O'Rourke,K., Boguski,M.S., & Dixit,V.M. A novel RING finger protein interacts with the cytoplasmic domain of CD40. J. Biol. Chem. 269, 30069-30072 (1994).

(98) Lee,E.G. et al. Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. Science 289, 2350-2354 (2000).

(99) Song,H.Y., Rothe,M., & Goeddel,D.V. The tumor necrosis factor-inducible zinc finger protein A20 interacts with TRAF1/TRAF2 and inhibits NF-kappaB activation. Proc. Natl. Acad. Sci. U. S. A 93, 6721-6725 (1996).

(100) Ishida,T. et al. CD40 signaling-mediated induction of Bcl-XL, Cdk4, and Cdk6. Implication of their cooperation in selective B cell growth. J. Immunol. 155, 5527-5535 (1995).

(101) Mauri,C., Mars,L.T., & Londei,M. Therapeutic activity of agonistic monoclonal antibodies against CD40 in a chronic autoimmune inflammatory process. Nat. Med. 6, 673-679 (2000).

(102) Kyburz,D. et al. Human rheumatoid factor production is dependent on CD40 signaling and autoantigen. J. Immunol. 163, 3116-3122 (1999).

(103) Gregersen,P.K. et al. REL, encoding a member of the NF-kappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. Nat. Genet.(2009).

(104) Hayden,M.S., West,A.P., & Ghosh,S. NF-kappaB and the immune response. Oncogene 25, 6758-6780 (2006).

(105) Zhou,H.J. et al. Nuclear CD40 interacts with c-Rel and enhances proliferation in aggressive B-cell lymphoma. Blood 110, 2121-2127 (2007).

(106) Croft,M. Co-stimulatory members of the TNFR family: keys to effective T-cell immunity? Nat. Rev. Immunol. 3, 609-620 (2003).

(107) Itoh,T., Ijuin,T., & Takenawa,T. A novel phosphatidylinositol-5-phosphate 4-kinase (phosphatidylinositol-phosphate kinase IIgamma) is phosphorylated in the endoplasmic reticulum in response to mitogenic signals. J. Biol. Chem. 273, 20292-20299 (1998).

(108) Castellino,A.M., Parker,G.J., Boronenkov,I.V., Anderson,R.A., & Chao,M.V. A novel interaction between the juxtamembrane region of the p55 tumor necrosis factor receptor and phosphatidylinositol-4-phosphate 5-kinase. J. Biol. Chem. 272, 5861-5870 (1997).

(109) Auger,I. et al. New autoantigens in rheumatoid arthritis: screening 8268 protein arrays with RA patients' sera. Ann. Rheum. Dis.(2008).

(110) Frazer,K.A. et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 449, 851-861 (2007).

(111) A haplotype map of the human genome. Nature 437, 1299-1320 (2005).

(112) Lee-Kirsch,M.A. et al. Mutations in the gene encoding the 3'-5' DNA exonuclease TREX1 are associated with systemic lupus erythematosus. Nat. Genet. 39, 1065-1067 (2007).

(113) Hugot,J.P. et al. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411, 599-603 (2001).

(114) Ogura,Y. et al. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 411, 603-606 (2001).

(115) Duerr,R.H. et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314, 1461-1463 (2006).

(116) Kotowski,I.K. et al. A spectrum of PCSK9 alleles contributes to plasma levels of low-density lipoprotein cholesterol. Am. J. Hum. Genet. 78, 410-422 (2006).

(117) Romeo,S. et al. Population-based resequencing of ANGPTL4 uncovers variations that reduce triglycerides and increase HDL. Nat. Genet. 39, 513-516 (2007).

(118) Cohen,J.C. et al. Multiple rare variants in NPC1L1 associated with reduced sterol absorption and plasma low-density lipoprotein levels. Proc. Natl. Acad. Sci. U. S. A 103, 1810-1815 (2006).

(119) Cohen,J.C. et al. Multiple rare alleles contribute to low plasma levels of HDL cholesterol. Science 305, 869-872 (2004).

(120) Ji,W. et al. Rare independent mutations in renal salt handling genes contribute to blood pressure variation. Nat. Genet. 40, 592-599 (2008).

(121) Cohen,J.C., Boerwinkle,E., Mosley,T.H., Jr., & Hobbs,H.H. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N. Engl. J. Med. 354, 1264-1272 (2006).

(122) Cohen,J. et al. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. Nat. Genet. 37, 161-165 (2005).

(123) Kathiresan,S. et al. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. Nat. Genet. 40, 189-197 (2008).

(124) Church,G.M. & Gilbert,W. Genomic sequencing. Proc. Natl. Acad. Sci. U. S. A 81, 1991-1995 (1984).

(125) Church,G.M. & Kieffer-Higgins,S. Multiplex DNA sequencing. Science 240, 185-188 (1988).

(126) Porreca,G.J. et al. Multiplex amplification of large sets of human exons. Nat. Methods 4, 931-936 (2007).

(127) Frazer,K.A., Murray,S.S., Schork,N.J., & Topol,E.J. Human genetic variation and its contribution to complex traits. Nat. Rev. Genet. 10, 241-251 (2009).

(128) Stefansson,H. et al. Large recurrent microdeletions associated with schizophrenia. Nature 455, 232-236 (2008).

(129) Rare chromosomal deletions and duplications increase risk of schizophrenia. Nature 455, 237-241 (2008).

(130) Walsh,T. et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. Science 320, 539-543 (2008).

(131) McCarroll,S.A. Extending genome-wide association studies to copy-number variation. Hum. Mol. Genet. 17, R135-R142 (2008).

(132) Thabet,M.M. et al. The contribution of Fc gamma receptor IIIA gene 158V/F polymorphism and copy number variation to the risk of ACPA positive rheumatoid arthritis. Ann. Rheum. Dis.(2008).

(133) Schaschl,H., Aitman,T.J., & Vyse,T.J. Copy number variation in the human genome and its implication in autoimmunity. Clin. Exp. Immunol. 156, 12-16 (2009).

(134) Cookson,W., Liang,L., Abecasis,G., Moffatt,M., & Lathrop,M. Mapping complex disease traits with global gene expression. Nat. Rev. Genet. 10, 184-194 (2009).

(135) Mattick,J.S. The genetic signatures of noncoding RNAs. PLoS. Genet. 5, e1000459 (2009).

(136) Smyth,D.J. et al. Shared and distinct genetic variants in type 1 diabetes and celiac disease. N. Engl. J. Med. 359, 2767-2777 (2008).

Nederlandse Samenvatting

Reumatoïde artritis (RA) is een chronische auto-immuunziekte met een prevalentie van één procent onder de Nederlandse bevolking. Ondanks dat de etiologie van RA nog grotendeels onbekend is, is bekend dat zowel erfelijke als genetische factoren betrokken zijn bij ontstaan en progressie van ziekte. Het onderzoek beschreven in dit proefschrift heeft zich geconcentreerd op kandidaat genen die mogelijk een rol spelen bij RA; dit om meer inzicht te krijgen in de erfelijke factoren die ten grondslag liggen aan het ziekteproces. Het identificeren van deze risicofactoren zal uiteindelijk leiden tot een beter begrip van het ziekteproces, en kan daardoor bijdragen tot een betere behandeling. Genetische risicofactoren worden gevonden doordat de natuurlijk voorkomende genetische variatie in een andere frequentie voorkomt in patiënten dan in personen zonder RA.

Het eerste deel van dit proefschrift (**Hoofdstukken 2-5**) beschrijft de ontdeking van een genetische risicofactor voor RA op chromosoom 9. Deze variatie ligt op een stukje DNA, dat voor twee genen codeert: *Tumour Necrosis Factor Receptor Associated Factor 1* (*TRAF1*) en *Complement factor 5* (*C5*). Deze ontdekking is inmiddels bevestigd in duizenden patiënten uit Nederland, Zweden, de VS, Frankrijk en Engeland. Echter, dit is een van de weinige risicofactoren waarvan ook in 1356 familieleden uit 452 families is aangetoond dat hij preferentieel met RA overerft. Deze variatie is niet alleen geassocieerd met RA maar ook met juvenile arthritis (JIA) en systemic lupus erythematosus (SLE). Dit toont aan dat deze risicofactor mogelijk in een biologische route ligt die een bijdrage levert aan deze verschillende auto-immuunziekten.

In het tweede deel van dit proefschrift wordt onderzoek beschreven over het interleukine-10 (IL10) gen. IL10 is een belangrijk ontstekingonderdrukkend cytokine en is bij vele aspecten van de immuunreactie betrokken. Uit voorafgaand onderzoek is gebleken dat DNA varianten in dit gen mogelijk verantwoordelijk zijn voor de verschillen in het maken van IL10 tussen mensen. Een lage IL10 produktie is geassocieerd met het ontwikkelen van familiale osteoartritis. In het kader hiervan heeft dit onderzoek het verband tussen genetische variaties van IL10 en osteoartritis bekeken. De bevindingen laten geen associatie tussen IL-10 en deze ziekte zien (**Hoofdstuk 6**).

Omdat IL10 zo'n belangrijke rol speelt bij het immuunsysteem is het onderzoek naar de invloed van ons DNA op het maken van dit cytokine zeer interessant. Wij hebben het gebied op het DNA waar de erfelijke variaties, geassocieerd met verschillen in IL10 produktie, gelegen kunnen zijn, nauwkeurig in kaart gebracht. Tevens hebben we aangetoond dat een haplotype in dit gebied gecorreleerd is met allel-specifieke transcriptie van het gen (**Hoofdstuk 7**). In **hoofdstuk 8** zijn de resultaten beschreven over een associatie van drie polymorfismen van het IL10 gen die mogelijk betrokken zijn bij de ontwikkeling van restenose. Restenose is het opnieuw vernauwen van een ader, nadat deze met een klein ballonnetje op een katheter is opgerekt om een verstopping op te heffen. IL10 is een interessante kandidaat voor verbetering van anti-restenose therapie, aangezien dit een van de meest bestudeerde anti-inflammatoire cytokinen is. IL10 is een risico marker voor het ontwikkelen van restenose. Een therapie gericht op IL-10 zou mogelijk kunnen bijdragen tot het verbeteren van individuele therapie, aangezien de cytokinen zoals IL-10 kunnen worden aangepakt door middel van een drug-eluting stent. In **Hoofdstuk 9** wordt de rol van genetische varianten in het IL10 gen in relatie tot RA beschreven. Dit lijkt geen belangrijke rol te spelen voor respectievelijk de ernst en de remissie van RA patienten.

In het derde deel van dit proefschrift wordt onderzoek gedaan naar de rol van verschillende andere genetische risicofactoren. In **Hoofdstuk 10** wordt de associatie tussen *TNFα* en de ernst van RA beschreven. Ondanks het feit dat RA patienten vaak grote hoeveelheden *TNFα* eiwitten maken en dat het gebruik van anti-TNF therapie in de kliniek effectief is, hebben wij geen correlatie tussen genetische varianten in dit gen en het verloop van de ziekte kunnen aantonen. Recentelijk is er een gen-gebied geïdentificeerd dat een TNF-gerelateerd molecuul in kaart brengt. In **Hoofdstuk 11**, wordt de link tussen *TNFα-induced protein 3 (TNFAIP3)* en de ernst van RA beschreven. Onze studie laat een correlatie zien tussen *TNFAIP3* en het krijgen van een ernstigere vorm van de ziekte.

In **Hoofdstuk 12** wordt de onafhankelijk replicatie in de Nederlandse populatie en een meta-analyse van drie beschreven genetische risico factoren (*STAT4*, *CTLA4* en *IL2/21*) bediscusciëerd. Wij laten zien dat *CTLA4* inderdaad geassocieerd is met RA maar dat deze associatie alleen gevonden wordt in een deel van de patiëntenpopulatie; de autoantilichaam positieve patiënten. In **Hoofdstuk 13**, wordt de bevinding dat het *IL2/21* gebied niet alleen een rol speelt bij RA, maar ook bij JIA, beschreven. In **Hoofdstuk 14**, worden genetische varianten in het *IRF5* gen beschreven. Deze lijken voornamelijk geassocieerd met het krijgen van een autoantilichaam-negatieve vorm van RA. In **Hoofdstuk 15** worden zeer recent gevonden genetische risco-factoren beschreven. Deze studie is verricht in het kader van een grote consortium studie, waarbij een significante associatie met een polymorfisme in CD40 is gevonden. Omdat CD40 een centrale rol vervult binnen het immuunsysteem, is het hoogst waarschijnlijk dat het gevonden gen belangrijk is voor de ontwikkeling en instandhouding van RA. In **Hoofdstuk 16** wordt de replicatie van twee belangrijke genen in de IL-2 route (IL2RA en IL2RB) beschreven. Deze route is nauw betrokken bij de totstandkoming van specifieke afweerreacties tegen antigenen. Deze studie geeft aan dat naast de CD40 signaleringscascade, de IL2-receptor signaleringscascade ook een belangrijk rol zou kunen spelen bij het ontwikkeling van RA.

De resultaten die worden beschreven in dit proefschrift hebben geleid tot nieuwe inzichten in de rol van vaak-voorkomende genetische varianten in een complexe ziekte als RA. Deze bevindingen zullen zeker leiden tot een beter begrip van het ziekteproces, wat in de toekomst mogelijk toegepast kan worden bij het ontwikkelen van nieuwe behandelingen voor RA.

Acknowledgements

I would like to thank everyone who, in one way or another, contributed to the birth and completion of this thesis.

To start with, I would like to thank Tom Huizinga and Rene Toes for being my mentors. Tom, your never-ending enthusiasm and your confidence in me have always been and remain a great source of support for me. Thank you!. Rene, your critical approach to science will never cease to amaze me. I have learned and continue to learn from you about the greatness of research. Thank you for being such a continuous source of inspiration!

The years I have spent at the Department of Rheumatology in Leiden have altogether been great years. Like most PhD students, I have faced my fair share of ups and downs but what remains of this experience is the tremendous wealth of knowledge I have gained, both in scientific and life experiences. For this great adventure, I would like to thank all of you. Joris, Nivine, Bep and Elsbet, you were a great source of moral support for me during my first year in Holland! For that I would like to extend a very special thank you! Joris, I am now finally convinced that you are not "German"☺. Gerrie, Ellen, Joris and Aleida, thank you for your tremendous help in the lab and for all the great conversations we had over coffee! Nina and Rute, I have thoroughly enjoyed working together with you in the last couple of years! Leonie, Wanda, Judith and Roelof, I truly enjoyed our discussions about life as a PhD student and I have also greatly valued your help in keeping my immunology going! Wang and Leendert, thank you for providing fresh perspective in the female-skewed room we had upstairs….. It often resulted in funny and enjoyable discussions! Kirsten, I will never forget the never-ending weeks we spent on Project X :-) Onno, it never ceased to amaze me how you juggled having so many students at the same time! It was always a fun experience to see you together with your students in the lab! Diane, it has been and remains great working together with you, hopefully I will one day make it to running night! Jeanine, thank you, your help with statistics were pivotal to many of the manuscripts in this thesis! Heleen, you were such a pleasure to work with. Thank you for the many fruitful discussions on juvenile arthritis, I have learned a lot! To all of my colleagues and friends in Leiden, I have valued every interaction, and they will forever form a part of my life!

I would also like to thank the people who have significantly contributed to many pleasant and fruitful collaborations that have resulted in some of the published work in this thesis: My colleagues at the Molecular Epidemiology department and the Leiden Genome Technology Center, Colleagues at the Medical Statistics department, Colleagues in the department of Complex Diseases in Utrecht, Colleagues at the Karolinska Institute in Sweden, Celera in the US and Crete in Greece.

I would also like to extend a very special thank you note to Andreea. Throughout the years you have become one of my best friends! You were always (and continue to be) there through the good and bad times! My years in Leiden would never have been the same without your true friendship!

To my present colleagues in Boston I thank you for the excellent environment. Robert, thank you for trusting me and providing me the opportunity to finish my thesis! Eli, Paul, Nikos, Linda and Lydia, although I have known you for a very short time, you have played a great part in my recent transition to Boston!

I am also indebted to my entire family in Mauritius for always believing in me. I would like to thank especially my parents for their endless wisdom as well as my brother and sister for their continuous support. Alis, having you in Holland has made a world of difference to my life! To my wonderful family, thank you for your unconditional love.

I would also like to thank Amri's parents; you have in the last years become a part of my extended family. I look forward to even more great adventures together!

Last but not least, Amri, words cannot express how wonderful, supportive and loving you have been. None of this would have been possible without you.

Fina

Curriculum Vitae

Name: Bibi Aisha Safinaz (Fina) Kurreeman
Date of Birth: September 22, 1979
Place of Birth: Plaines Wilhems, Mauritius

1996-1998	Queen Elisabeth College, **Mauritius** High School Graduate, Cambridge Board 'A' levels.
1998-2001	University of Westminster, London, **UK** Bsc (Hons) Biotechnology in Medicine
2001-2003	Leiden University, Leiden, **Netherlands** Msc Biomedical Sciences Master thesis Supervisor: Prof. Dr .T.W.J Huizinga
2003-2008	Leiden University Medical Center, Leiden, **Netherlands** Ph.D. at the Department of Rheumatology Promoter: Prof.Dr.T.W.J.Huizinga Co-promoter: Dr. R.E.M. Toes
2008-2009	Leiden University Medical Center, Leiden, **Netherlands** Post-Doctoral Fellow at the Department of Rheumatology Supervisors: Prof. Dr. T.W.J.Huizinga and Dr. Rene Toes
2009-	Brigham and Women's Hospital, Harvard Medical School, **USA** Post-Doctoral Fellow Principal Investigator: Dr. R.M.Plenge

List of Publications

Kurreeman, F.A., Daha, N.A., Chang, M., Catanese, J.J., Begovich, A.B., Huizinga, T.W.J.,Toes, R.E.M. Association of IL2RA and IL2RB with rheumatoid arthritis - a replication study in a Dutch population. Ann.Rheum.Dis In press.

van der Linden,M.P, Feitsma, A.L, le Cessie, S, Kern, M, Olsson, L.M, Raychaudhuri, S, Begovich, A.B, Chang, M, Catanese, J.J, **Kurreeman**, **F.A**, van Nies, J, van der Heijde, D.M, Gregersen, P.K, Huizinga, T.W, Toes, R.E, van der Helm-Van Mil, A.H. **2009**. Association of a single-nucleotide polymorphism in CD40 with the rate of joint destruction in rheumatoid arthritis. *Arthritis Rheum*. 60(8): 2242-7**.**

Keijser, S, **Kurreeman, F.A**, de Keizer, R.J, Dogterom-Ballering, H, van der Lelij, A, Jager, M.J, Nibbering, P.H. **2009.** IL-10 promotor haplotypes associated with susceptibility to and severity of bacterial corneal ulcers. *Exp Eye Res*. 88(6):1124-8.

Kurreeman, F.A., Goulielmos, G.N., Alizadeh, B.Z., Rueda, B., Houwing-Duistermaat, J., Sanchez, E., Bevova, M., Radstake, T.R., Vonk, M.C., Galanakis, E., Ortego, N., Verduyn, W., Zervou, M.I., SLEGEN Consortium, Roep, B.O., Dema, B., Espino, L., Urcelay, E., Boumpas, D. T., van den Berg, L.H., Wijmenga, C., Koeleman, B.P., Huizinga, T.W., Toes, R.E., & Martin, J. **2009**. The TRAF1-C5 region on chromosome 9q33 is associated with multiple autoimmune diseases. *Ann.Rheum.Dis.*.

Kurreeman, F. A., Daha N.A, Marques, R. B., Stoeken-Rijsbergen, G., Verduijn, W., Huizinga, T. W., & Toes, R. E. **2009**. Confirmation of STAT4, IL2/IL21, and CTLA4 polymorphisms in rheumatoid arthritis. *Arthritis Rheum*. 60(5): 1255-1260.

Scherer, H. U., van der Linden, M. P., **Kurreeman, F. A**., Stoeken-Rijsbergen, G., le, C. S., Huizinga, T. W., van der Helm-van Mil AH, & Toes, R. E. **2009**. Association of the 6q23 region with the rate of joint destruction in rheumatoid arthritis. *Ann.Rheum.Dis.*.

Albers, H. M., **Kurreeman, F. A**., Stoeken-Rijsbergen, G., Brinkman, D. M., Kamphuis, S. S., van Rossum, M. A., Girschick, H. J., Wouters, C., Saurenmann, R. K., Hoppenreijs, E., Slagboom, P., Houwing-Duistermaat, J. J., Verduijn, W., Huizinga, T. W., Ten, C. R., Toes, R. E., & Schilham, M. W. **2009**. Association of the autoimmunity locus 4q27 with juvenile idiopathic arthritis. *Arthritis Rheum.*, 60(3): 901-904.

Thabet, M. M., Huizinga, T. W., Marques, R. B., Stoeken-Rijsbergen, G., Bakker, A. M., **Kurreeman, F. A.**, White, S., Toes, R. E., & van der Helm-van Mil AH **2008**. The contribution of Fc gamma receptor IIIA gene 158V/F polymorphism and copy number variation to the risk of ACPA positive rheumatoid arthritis. *Ann.Rheum.Dis.*.

Raychaudhuri, S., Remmers, E. F., Lee, A. T., Hackett, R., Guiducci, C., Burtt, N. P., Gianniny, L., Korman, B. D., Padyukov, L., **Kurreeman, F. A.**, Chang, M., Catanese, J. J., Ding, B., Wong, S., van der Helm-van Mil AH, Neale, B. M., Coblyn, J., Cui, J., Tak, P. P., Wolbink, G. J., Crusius, J. B., van der Horst-Bruinsma IE, Criswell, L. A., Amos, C. I., Seldin, M. F., Kastner, D. L., Ardlie, K. G., Alfredsson, L., Costenbader, K. H., Altshuler, D., Huizinga, T. W., Shadick, N. A., Weinblatt, M. E., de, V. N., Worthington, J., Seielstad, M., Toes, R. E., Karlson, E. W., Begovich, A. B., Klareskog, L., Gregersen, P. K., Daly, M. J., & Plenge, R. M. **2008**. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. *Nat.Genet.*, 40(10): 1216-1223.

Kurreeman, F. A., Rocha, D., Houwing-Duistermaat, J., Vrijmoet, S., Teixeira, V. H., Migliorini, P., Balsa, A., Westhovens, R., Barrera, P., Alves, H., Vaz, C., Fernandes, M., Pascual-Salcedo,

D., Michou, L., Bombardieri, S., Radstake, T., van, R. P., van de, P. L., Lopes-Vaz, A., Prum, B., Bardin, T., Gut, I., Cornelis, F., Huizinga, T. W., Petit-Teixeira, E., & Toes, R. E. **2008**. Replication of the tumor necrosis factor receptor-associated factor 1/complement component 5 region as a susceptibility locus for rheumatoid arthritis in a European family-based study. *Arthritis Rheum.*, 58(9): 2670-2674.

Chang, M., Rowland, C. M., Garcia, V. E., Schrodi, S. J., Catanese, J. J., van der Helm-van Mil AH, Ardlie, K. G., Amos, C. I., Criswell, L. A., Kastner, D. L., Gregersen, P. K., **Kurreeman, F. A.**, Toes, R. E., Huizinga, T. W., Seldin, M. F., & Begovich, A. B. **2008**. A large-scale rheumatoid arthritis genetic study identifies association at chromosome 9q33.2. *PLoS.Genet.*, 4(6): e1000107.

Kurreeman, F. A., Albers, H. M., Houwing-Duistermaat, J. J., Brinkman, D. M., Kamphuis, S. S., Girschick, H. J., Wouters, C., van Rossum, M. A., Verduijn, W., Toes, R. E., Huizinga, T. W., Schilham, M. W., & Ten, C. R. **2008**. The TRAF1/C5 region is a risk factor for polyarthritis in juvenile idiopathic arthritis. *Ann.Rheum.Dis.*, 67(11): 1578-1580.

Kurreeman, F. A., Padyukov, L., Marques, R. B., Schrodi, S. J., Seddighzadeh, M., Stoeken-Rijsbergen, G., van der Helm-van Mil AH, Allaart, C. F., Verduyn, W., Houwing-Duistermaat, J., Alfredsson, L., Begovich, A. B., Klareskog, L., Huizinga, T. W., & Toes, R. E. **2007**. A candidate gene approach identifies the TRAF1/C5 region as a risk factor for rheumatoid arthritis. *PLoS.Med.*, 4(9): e278.

Sigurdsson, S., Padyukov, L., **Kurreeman, F. A**., Liljedahl, U., Wiman, A. C., Alfredsson, L., Toes, R., Ronnelid, J., Klareskog, L., Huizinga, T. W., Alm, G., Syvanen, A. C., & Ronnblom, L. **2007**. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. *Arthritis Rheum.*, 56(7): 2202-2210.

van der Helm-van Mil AH, **Kurreeman, F. A.**, Toes, R. E., & Huizinga, T. W. **2007**. Association of tumor necrosis factor alpha polymorphism and radiographic progression in rheumatoid arthritis: comment on the article by Khanna et al. *Arthritis Rheum.*, 56(3): 1032-1033.

Monraats, P. S., **Kurreeman, F. A.**, Pons, D., Sewgobind, V. D., de Vries, F. R., Zwinderman, A. H., de Maat, M. P., Doevendans, P. A., de Winter, R. J., Tio, R. A., Waltenberger, J., Huizinga, T. W., Eefting, D., Quax, P. H., Frants, R. R., van der, L. A., van der Wall, E. E., & Jukema, J. W. **2007**. Interleukin 10: a new risk marker for the development of restenosis after percutaneous coronary intervention. *Genes Immun.*, 8(1): 44-50.

Riyazi, N., **Kurreeman, F. A.**, Huizinga, T. W., Dekker, F. W., Stoeken-Rijsbergen, G., & Kloppenburg, M. **2005**. The role of interleukin 10 promoter polymorphisms in the susceptibility of distal interphalangeal osteoarthritis. *J.Rheumatol.*, 32(8): 1571-1575.

Kurreeman, F. A., Schonkeren, J. J., Heijmans, B. T., Toes, R. E., & Huizinga, T. W. **2004**. Transcription of the IL10 gene reveals allele-specific regulation at the mRNA level. *Hum.Mol.Genet.*, 13(16): 1755-1762.