

New insights into genetics of immune responses in rheumatoid arthritis

A. Ruysen-Witrand^{1,2,3}, A. Constantin^{1,2,3}, A. Cambon-Thomsen^{1,2}, M. Thomsen^{1,2}

1 Inserm, UMR1027, Toulouse, France

2 University of Toulouse III - Paul Sabatier, UMR1027, Toulouse, France

3 CHU Toulouse Purpan, Rheumatology Center, Toulouse, France

Corresponding Author:

Mogens Thomsen, Inserm UMR 1027, Faculty of Medicine,

37 allées Jules Guesde, F-31000 Toulouse

E-mail: thomsen@cict.fr

Short title: Genetics of rheumatoid arthritis

Abstract

Rheumatoid Arthritis (RA) is a common autoimmune disease with a strong genetic component. Numerous aberrant immune responses have been described during the evolution of the disease. In later years the appearance of anti-citrullinated protein antibodies (ACPAs) has become a hallmark for the diagnosis and prognosis of RA. The posttranslational transformation of arginine residues of proteins and peptides into citrulline (citrullination) is a natural process in the body, but for unknown reasons autoreactivity towards citrullinated residues may develop in disposed individuals. ACPAs are often found years before clinical manifestations. ACPAs are present in about 70% of RA patients and constitute an important disease marker, distinguishing patient groups with different prognosis and different response to various treatments. Inside the HLA region, some *HLA-DRB1* alleles are strongly associated with their production. Genome-wide association studies in large patient materials have defined a great number of single nucleotide polymorphisms (SNPs) outside of the HLA region that are associated with ACPA positive RA. The SNPs are generally located close to or within genes involved in the immuneresponse or signal transduction in immune cells. Some environmental factors such as tobacco smoking are also positively correlated with ACPA production. In this review we will attempt to describe the genes and loci associated with ACPA positive RA or ACPA negative RA and to clarify their potential role in the development of the disease.

Keywords:

Rheumatoid arthritis, single nucleotide polymorphisms, genetics, immune response, autoantibodies, citrullination

Introduction

Autoimmune diseases are phenotypically heterogeneous. Some are organ-specific such as type-1 diabetes, myasthenia gravis and Graves' disease, and others are systemic as Systemic Lupus Erythematosus (SLE), Sjögren's disease or Rheumatoid Arthritis (RA). They are all multifactorial diseases, with a complex interplay between genes and environment. The principal susceptibility genes are located in the major histocompatibility complex, the HLA region in humans. In RA patients, the susceptibility is primarily associated with certain *HLA-DRB1* alleles (1). Most recently, associations with *HLA-B* and *HLA-DP* have been demonstrated when combining 6 large independent genome-wide datasets (2). Genome-wide association studies (GWAS) using single nucleotide polymorphisms (SNPs) have indicated that in addition to genes in the HLA region, a multitude of other genes may be involved in disease susceptibility, in particular genes encoding immune regulatory factors (3). Among environmental factors, tobacco smoking has been shown to be associated with RA (4).

Rheumatoid arthritis is the most common chronic inflammatory arthritis affecting 0.5 to 1% of adults in developed countries (5). Three times more women than men are affected and the mean age-of-onset is about 50 years (6). The presence of anti-citrullinated protein antibodies (ACPAs) has become an important diagnostic and prognostic criterion for RA, and it has been suggested that ACPA positivity or negativity may define different disease entities (7, 8). In order to take into account the importance of ACPAs for the diagnosis of RA, new classification criteria for RA were established in 2010 (9). In this review the possible role of candidate genes for the development of ACPA positive (ACPA+) RA will be discussed. The overview represents an update of a review in French published previously by one of us (10), extended to take into account recent GWAS data and to integrate the present knowledge on genes involved.

Pathophysiology

Targets of autoantibodies

The presence of autoantibodies, in particular rheumatoid factor (RF) and anti-citrullinated peptide antibody (ACPA), is a well-established indicator of RA disease severity. The antibodies may precede the onset of the disease by many years (11-13). RF is a classical pathogenic marker consisting of IgM and IgA antibodies directed against the Fc fragment of IgG. ACPAs seem to be more specific and sensitive for diagnosis of RA and are also better predictors of poor prognostic features such as progressive

joint destruction. Although ACPAs have a good specificity for RA diagnosis, about 95%, the sensitivity is somewhat lower, 67% (14). The corresponding figures for RF are 85% and 69%, respectively. Clinical studies show that RA patients with both RF and ACPAs differ from the so-called autoantibody negative patients in term of disease activity, structural severity and response to various treatments(12, 15).

Citrullination

Citrullination of proteins is a post-transcriptional modification featuring the conversion of a positively loaded arginine residue into the neutral citrulline, see figure 1. This conversion may alter the three-dimensional structure of proteins and also generate neo-epitopes on peptides presented by HLA molecules. Citrullination of proteins is a constitutive phenomenon in many tissues, in particular observed in the normal epidermis. It is caused by an enzyme called peptidyl-arginine deiminase (PAD) with 5 isoforms expressed differently in various tissues(16). Citrullination may also be induced under inflammatory conditions in the synovium of patients with rheumatoid arthritis or in the upper airways of smokers (17). Among the proteins in the joint that undergo citrullination can be mentioned fibrin, vimentin, fibronectin and alpha enolase(18). Studies in experimental animals show that citrullinated peptides are generated constitutively in dendritic cells and macrophages, and in B cells after engagement of the antigen receptor(19). Citrullinated peptides fit in general better into HLA-DR molecules because charged arginine residues are replaced by neutral citrulline residues. In disposed individuals a loss of immune tolerance to citrullinated proteins and peptides may develop because T-cells in the periphery are confronted with peptide-HLA complexes against which they have not been tolerized in the thymus. T cells from ACPA positive RA patients recognize citrullinated vimentin peptides *in vitro* restricted by HLA-DR4 and similar findings have been made in *HLA-DR4* transgenic mice(20). It has been shown that peripheral B lymphocytes from a majority of RA patients are able to produce ACPA after non-specific stimulation *in vitro*(21). The production of antibodies was mainly detected in cultures from patients with HLA-DR alleles associated with RA. Interestingly, B cells from about one third of normal subjects carrying these HLA-DR alleles, but without family history of RA, also produced ACPA antibodies *in vitro*.

Dysregulation of the Immune response:

The presence of autoantibodies years before clinical manifestations points towards an important role of the adaptive immune system in the early pathogenesis of RA, at least in the ACPA+ patients(22). T

cells are found abundantly in inflamed joints and autoreactive T cells have been identified within the joints (23). RA is traditionally considered to be mediated by type 1 helper T cells, but the role of so-called type 17 helper T cells is increasingly evoked (24). A great number of cytokines are produced by such cells and TNF- α together with IL-17A provoke activation of synoviocytes and chondrocytes (25). Humoral adaptive immunity is central in RA (26). Synovial B lymphocytes are localized in aggregates together with T cells and in some tissues can be found ectopic follicles. The efficacy of clinical treatment of antibodies towards CD20 suggests that their role is not limited to the production of autoantibodies but also include production of certain cytokines (27). A great number of innate effector cells are found in the synovial membrane and/or synovial fluid, and macrophages are important as they release a multitude of inflammatory cytokines (28). ACPAs have been shown to enhance NF- κ B activity and TNF- α production in monocytes and macrophages by binding citrullinated GRp78, expressed on the surface of these cells(29). Neutrophils and mast cells also contribute to the synovitis by production of proteases and other inflammatory intermediates (30, 31). Many cytokines contribute to the disease process and TNF- α plays an important role as mentioned above. Other important cytokines are IL-6 and the IL-1 cytokine family (32) and much research has been dedicated to unraveling the intracellular signaling pathways that regulate cytokine production (33, 34). The synoviocytes also participate in the disease process as they develop a phenotype that contributes to local cartilage destruction and sustains T and B cell survival (35, 36). Figure 3 shows schematically the initiation of the immune response in RA and the inflammatory phase of the disease.

Genetic factors associated with ACPA+ RA

Most of the factors described below and in table 1 have been defined in genome wide association studies (GWAS). A metaanalysis of results from independent GWAS have permitted to accumulate very large materials. Stahl and colleagues from Europe and North America have for instance first done a metaanalysis on 5,839 autoantibody positive RA patients and 20,169 controls followed by replication in an independent set of 6,768 patients and 8,806 controls, all of European descent (37). Due to the large number of comparisons made, an association is only considered significant when p is less than 5×10^{-8} . When the SNPs at risk are not located within a gene, the so-called GRAIL analysis has been performed to assign the most probable candidate gene(s) in linkage disequilibrium with the SNP (38). Kurreeman and colleagues have recently used electronic health records to track RA patients and obtain biospecimens for SNP analysis and antibody testing(39). In order to ensure European ancestry, a great number of ancestry informative markers were used. This also permitted to define

patients and controls of other ethnic origins (e.g. Asians, Africans and Hispanics). A number of those results are also reported in table 1.

HLA Genes

The association between HLA and RA was reported by several groups in the mid-seventies. A rather weak association with HLA-B was first reported, and when cellular and serological methods for defining HLA class II antigens became available, the association with HLA-DR4 was found to be stronger (40, 41). The possible associations with HLA-B were then considered to be due to linkage disequilibrium between *HLA-B* and *HLA-DR* alleles. The development of molecular biology methods revealed that a number of *HLA-DRB1* alleles were associated with the disease and it was then proposed that a shared epitope (SE) of the DR β 1 chain (amino acid residues 70-74) was the main factor causing the association between *HLA-DRB1* and RA(1). These residues are located on the α -helix of the DR β 1 chain. In a very recent analysis (GWAS) on a large combined material of European origin, amino acid polymorphisms were tested across the HLA class I and II antigens (2). As shown in table 1, four amino acid positions of HLA-DR β were important for the susceptibility to ACPA+ RA, namely positions 11, 13, 71 and 74. The two latter are part of the so-called shared epitope on the α helix (1) while the two others are located in the floor of the peptide binding groove, as shown in Figure 2. All 4 amino acids have side-chains pointing into the groove and are thus important for the peptide binding properties of the HLA molecule. On the basis of the polymorphisms of amino acids in these four positions, the DRB1 alleles can be classified according to their association with ACPA+ RA. *DRB1*04:01* is the most strongly associated, with an odds ratio of 4.44 in Europeans. The same authors showed an independent association for *HLA-B*08* that encodes an asparagine in position 9 and for a group of *HLA-DP1* alleles that encode a phenylalanine in position 9 on the β chain. For both molecules, position 9 is located in the floor of the peptide binding groove. A recent study in ethnically mixed RA patients and controls suggested that also position 67 on the DR β 1 chain was associated with the disease(42). It could be shown that this position is important for the binding of a citrullinated vimentin peptide, a likely auto-antigen in RA. In non-European populations, additional *HLA-DR* alleles may be associated. A recent study in Japanese showed no HLA association for ACPA- RA, and the strongest association for ACPA+ RA was observed for *DRB1*04:05*(43). Interestingly, the data also suggest significant and independent associations with other HLA loci than *HLA-DRB1*(*HLA-DP*, *HLA-B* and *HLA-C*).

Non HLA Genes

The genes below are listed according to their position on chromosomes.

PTPN22 (1p13)

The *PTPN22* gene encodes a tyrosine phosphatase (protein tyrosine phosphatase, non-receptor type 22), also called LYP. It negatively regulates activation of T and B cells via the T cell antigen receptor (TCR) and B cell antigen receptor (BCR). One SNP (rs2476601), characterized by a C/T mutation in position 1858, gives rise to the substitution of an arginine (R) by a tryptophan (W) in position 620 of the protein, in the binding domain of LYP. The 620W variant, encoded by allele 1858T, seems to be associated with a gain in catalytic function of LYP, resulting in a reinforcement of the negative regulation of activation of T and B cells. This may result in a reduced negative selection of auto reactive T and/or B precursors as well as in a reduced activity of regulatory T cells, thus promoting autoimmunity. Several independent studies have reported an association between carriage of allele 1858T and susceptibility to ACPA+ RA (44). During the preclinical phase of rheumatoid arthritis, carriage of allele 1858T is significantly associated with positivity of ACPAs (45). This allele is not found in Asian populations (46), but a polymorphism in the promoter region seems to be associated with RA in Chinese (47). A recent GWAS showed an association of another SNP located in the *PTPN22* gene, rs6679677, in a European sample (48). A weaker association was found with ACPA- RA.

CD2, CD58 (1p13.1)

CD2 is a cell adhesion molecule with costimulatory properties. It is present on the surface of peripheral T cells and some NK cells. The ligand is CD58, a member of the immunoglobulin superfamily, present on antigen presenting cells. The ligation of CD2 and CD58 causes activation of T lymphocytes and NK cells. The SNP rs11586238 is about 50 kb upstream of the *CD2* gene and also close to other important genes as *CD58* and *IGSF* (49).

TNFRSF14 (1p36)

The protein TNFRSF14 is member 14 of the TNF receptor superfamily. The cytoplasmic region of this receptor binds to several TRAF (TNF receptor associated factors) family members, which may mediate signal transduction pathways that activate the immune response. The SNP rs3890745 is close to the TNFRSF14 gene but is actually located in an intron of the *MMEL1* gene, which codes for a neutral endopeptidase. However, no polymorphism of this gene is suspected to be associated with RA and the GRAIL analysis pointed to TNFRSF14 as the causative gene (39).

Peptidyl arginine deiminase, type 4 (PADI4) (1p36.13)

PADI4 is member of a family of 5 calcium dependent enzymes that are expressed in various tissues, responsible for transformation of arginine residues into citrulline. PADI4 is found in the cytoplasm of monocytes, T and B cells, polynuclear neutrophils and eosinophils, as well as in natural killer cells. The PADI4 enzyme may translocate into the nucleus of activated cells and it plays a physiological role in regulating the transcriptional activity of numerous genes by counterbalancing the methylation of arginine residues of histones. A number of SNPs have been defined that characterize *PADI4* haplotypes influencing ACPA positivity in RA populations of Asian origin (46, 50). However, a metaanalysis of all European studies failed to show an association between *PADI4* haplotypes and RA (51).

FCGR (Fc Fragment of IgG Receptor) (1q21-23)

Receptors for the Fc fragment of IgG (FcγR) play an important role in the immune response. In humans, eight *FCGR* genes are assembled in the 1q21-23 region. FcγRIIb (CD32B) is generally considered to regulate negatively the activity of B cells. One non-synonymous SNP (rs1050501) in the *FCGR2B* gene causes the substitution of an isoleucine with a threonine in the transmembrane segment (187 Ile/Thr) of FcγRIIb, affecting its immunoregulatory functions. A study carried out in Taiwanese patients with RA showed an overrepresentation of allele 187-Ile in the ACPA+ RA patient group, in comparison with controls (52). FcγRIIIa (CD16) is generally considered to regulate positively the activity of monocytes/macrophages. One non-synonymous SNP (rs396991) in the gene *FCGR3A* causes the substitution of a valine with a phenylalanine in position 158 (158 V/F) of FcγRIIIa, affecting its immunoregulatory functions. A study carried out in Dutch patients with RA revealed an association between genotype VV and susceptibility to in ACPA+ RA but not of ACPA- RA (53). This study was confirmed by a British group that also performed a metaanalysis showing an association between VV genotype and ACPA + RA (54). Another metaanalysis showed an association between one SNP (rs12746613) located on 1q23.3 near *FCGR2A* genes and RA (49).

Deletion of the Late Cornified Envelope gene (LCE3C_LCE3B del) (1q21.3)

The epidermal differentiation complex contains several genes that are expressed in different types of epithelia such as skin, intestine and lung. An association was observed between patients homozygous for the deletion of the above mentioned gene and ACPA+ RA in two cohorts of RA (55). The same deletion is associated with psoriasis.

Protein tyrosine phosphatase, receptor type, C. PTPRC (1q31.3)

This enzyme is member of the protein tyrosine phosphatase (PTP) family, essential regulator of T- and B-cell antigen receptor signaling; it suppresses JAK kinases, and thus functions as a regulator of

cytokine receptor signaling. The SNP that has been shown to be associated with RA (rs10919563) is located within a *PTPRC* intron, 35 kb away from a rare non synonymous SNP that alters *PTPRC* splicing(49).

REL (2p13)

The protein c-Rel is a transcription factor that is a member of the Rel/NFκB family. The 'Rel domain,' is responsible for DNA binding, dimerization, nuclear localization, and binding to the NFκB inhibitor. C-Rel controls multiple steps in the development of regulatory T cells(56).

SPRED2 (2p14)

The Sprouty-related EVH1 domain containing protein regulates growth factor induced activation of the MAP kinase pathway in CD45+ hematopoietic cells(37).

AFF3 (2q11)

AFF3 is an RNA binding protein. It is a tissue restricted nuclear transcriptional activator preferentially expressed in lymphoid tissue(57).

STAT4 (2q32)

STAT4 is member of the STAT family of transcription factors which are essential for mediating responses to IL12 in lymphocytes, and regulating the differentiation of T helper cells(58).

CD28 (2q33)

CD28 is a co-stimulatory molecule required for T cell activation. Its ligands are CD80 and CD86 molecules on antigen presenting cells(49).

CTLA4 (2q33)

The CTLA-4 molecule belongs to an immunoglobulin superfamily expressed by activated T cells. CTLA-4 binds to CD80/CD86 molecules expressed by antigen-presenting cells. It negatively regulates activation of T cells by preventing the co-stimulatory CD28 molecules from binding to CD80/86 molecules (8). A study involving Dutch RA patients, together with a meta-analysis of published studies concluded that there was an association between a SNP (rs3087243) of gene CTLA4 and susceptibility to RA in ACPA+ Caucasoid patients, but not in ACPA- patients (59). A recent metaanalysis in populations of European descent confirmed the association (48).

PX domain containing serine/threonine kinase (PXX) (3p14)

Phlox homology (PX) domains bind to phosphoinositides and PDK is a serine/threonine kinase. It modulates Na-K ATPase enzymatic and ion pump activities. A SNP located in *PDK* gene (rs13315591) is associated with ACPA+ RA(48).

Recombination signal binding protein for immunoglobulin kappa J region (RBPJ) (4p15)

RBPJ is a transcriptional regulator important in the Notch signaling pathway that can bind specifically to the recombination signal sequence of immunoglobulin kappa type J segments. A significant association between a SNP located in *RBPJ* gene (rs874040) and ACPA+ RA was identified(48).

IL2, IL21 (4q27)

The IL2 cytokine plays an important role for T cell homeostasis and survival. IL21 has a central role for antibody mediated immune responses. It acts on NK cells, CD4+ T and B lymphocytes to induce antibody production and mediate class switching. It also induces Th17 differentiation and Th17 T cells are mediators of inflammation(60).

Ankyrin repeat domain 55; Interleukin 6 signal transducer (ANKRD55, IL6ST) (5q11)

The function of the *ANKRD55* gene is unknown. *IL6ST* is a signal transducer shared by many cytokines, including interleukin 6 (IL6). A significant association is present between a SNP (rs6859219) located near *IL6ST* gene and RA susceptibility (37). A significant association between a SNP located in *ANKRD55* gene (rs10040327) near *IL6ST* gene and ACPA+ RA was identified(48).

Beta-2-adrenergic receptor (ADRB2) (5q32-q34)

The *ADRB2* gene encodes the β_2 adrenergic receptor. The lymphocytes of patients with RA are characterized in particular by a lesser expression of β_2 adrenergic receptors, correlated negatively to the activity of the disease. In German RA patients, a polymorphism causing a non-synonymous mutation of codon 16 (Gly16Arg), characterized by the substitution of a glycine with an arginine, was found to be associated with susceptibility to the disease as well as with a positivity of ACPAs (ACPA+ in 93 % of homozygotes for the allele encoding arginine compared with 75 % of homozygotes for the allele encoding glycine) (61).

Allograft inflammatory factor 1 (AIF-1) (6p21.3)

Allograft inflammatory factor 1 (AIF-1) is a protein which plays a role in certain inflammatory and autoimmune processes. It is notably overexpressed by mononuclear cells and fibroblast-like synoviocytes in the joints of RA patients. The *AIF-1* gene is located in the MHC class III region. One study looked for an association between a non-synonymous (Arg15Trp) single nucleotide

polymorphism (rs2269475), located in intron 3 of isoform 2 of AIF-1 and RA in Polish patients. The T allele of this polymorphism proved to be associated with susceptibility to ACPA+ RA (62). The same group suggested recently that the SNP rs2259571 was associated with an active form of RA but not with ACPA positivity(63).

PR domain containing 1 (PRDM1) (6q21)

The protein acts as a repressor of beta-interferon gene expression. It binds specifically to a regulatory element of the IFN beta gene promoter. A positive association was observed between rs548234 near *PRDM1* and ACPA+ RA (49)

TNFAIP3 (6q23)

Genome-wide studies have identified genetic markers associated with susceptibility to RA in region 6q23 (64). Two SNPs (rs6920220 and rs10499194) were correlated to the status of the ACPAs(65). The markers are close to the gene encoding tumor necrosis factor alpha-induced protein 3 (TNFAIP3). TNFAIP3 is a negative regulator of NF- κ B and is involved in inhibiting TNF-receptor-mediated signaling effects.

T-cell activation RhoGTPase activating protein (TAGAP) (6q25)

The protein activates Rho GTPase during T cell activation. A significant association has been found between a SNP located in *TAGAP* gene (rs394581) and ACPA+ RA (48, 49).

Chemokine (C-C motif) receptor 6 (CCR6) (6q27)

Member of the beta chemokine receptor family expressed by immature dendritic cells and memory T cells, it is important for B-lineage maturation and antigen-driven B-cell differentiation, and it may regulate the migration and recruitment of dendritic cells and T cells during inflammatory and immunological responses. GWAS identified a significant association between a SNP (rs3093023) located in *CCR6* gene and RA susceptibility (37, 48).

Interferon regulatory factor 5 (IRF5) (7q32)

IRF5 belongs to a family of transcriptional factors which regulate, in particular, the system of type I interferons. As a ligand of several toll-like receptors (TLRs 4, 5, 7/8 and 9), IRF5 plays a central role in both innate and adaptive immunity, by activating, through nuclear factor kappa-B (NF- κ B), transcription of the genes of numerous pro-inflammatory cytokines such as TNF- α , IL-6 and IL-12p40. After an initial negative study published in 2006 (66), a second case-control study looked for an association between various SNPs located in gene *IRF5* and susceptibility to RA, in patients from

Swedish (1530 cases and 861 controls) and Dutch (387 cases and 181 controls) cohorts. This study revealed the promoting effect of a haplotype called H1 (frequent alleles of four SNPs) and the protector effect of a haplotype called H2 (rare alleles of four SNPs) with regard to the risk of RA in the Swedish cohort(67). The effect was more marked in ACPA- than in ACPA+ patients. Another case-control study, supplemented by a meta-analysis of case-control studies published in 2007, confirmed the existence of an association between various SNPs of the *IRF5* gene and protection with regard to the risk of RA, notably for one SNP (rs2004640). There too, this association was more marked in ACPA- patients than in ACPA+ patients (68). A more recent GWAS found an association between the T allele of a SNP located in *IRF5* (rs10488631) and ACPA+ RA, but not with ACPA-RA in a European population sample (48).

B lymphoid tyrosine kinase (BLK) (8p23)

This gene encodes a non-receptor tyrosine-kinase of the src family of proto-oncogenes that are typically involved in cell proliferation and differentiation. The protein has a role in B-cell receptor signaling and B-cell development. A significant association between a SNP located in *BLK* gene (rs13277113) and ACPA+ RA has been shown(48). This association was also confirmed in a GWAS performed in a Korean population sample of 801 cases of RA with either ACPA or RF antibodies versus 757 controls (46).

Chemokine (C-C motif) ligand 21 (CCL21) (9p13)

CCL21 is a cytokine that inhibits hemopoiesis and stimulates chemotaxis. It is chemotactic in vitro for thymocytes and activated T cells, and may also play a role in mediating homing of lymphocytes to secondary lymphoid organs. The SNP rs2812378 is weakly associated with the presence of ACPA in RA patients (37, 69).

Locus TRAF1-C5 (9q33-34.1)

The TNF receptor associated factor 1 (TRAF1) is a protein that associates with and mediates the signal transduction from various members of the TNF Receptor superfamily. C5 (fifth component of complement) genes located close to the TRAF1 gene. Two SNPs close to these genes, (rs3761847 and rs2900180), have shown to be positively associated with ACPA+ RA (65). Another SNP located in *TRAF1-C5* gene, rs10118357, was associated with ACPA+ RA (48).

IL2RA (10p15)

The interleukin 2 receptor alpha and beta chains, together with the common gamma chain, constitute the high-affinity IL2 receptor. A weak association with ACPA/RF positive RA has been reported(37).

PRKCQ (10p15)

Protein kinase C theta is one of the PKC family members. It is a calcium-independent and phospholipid-dependent protein kinase. This kinase is important for T-cell activation. It is required for the activation of the transcription factors NF-κB and AP-1(70).

Recombination activating gene 1/TNF receptor-associated factor 6, E3 ubiquitin protein ligase

RAG1/TRAF6 gene (11p12)

RAG1 is involved in activation of immunoglobulin V-D-J recombination. TRAF6 is member of the TNF receptor associated factor (TRAF) protein family;it mediates signaling from members of the TNF receptor superfamily as well as the Toll/IL-1 family.A positive association was observed between the SNP rs540386 within a *TRAF6* intron and RA (49).

IKZF3 (17q12)

IKZF3(IKAROS family zinc finger 3) participates in regulation and proliferation of B cells. Mice lacking the gene develop an SLE like syndrome. The SNP rs2872507 is associated with ACPA+ RA (39)

CD40 (20q13)

As a member of the TNF-receptor superfamily of costimulatory molecules, CD40 essential in mediating a broad variety of immune and inflammatory responses including T cell-dependent immunoglobulin class switching, memory B cell development, and germinal center formation. A significant association between a SNP located in *CD40* gene (rs4810485) and ACPA+ RA has been found (48, 69).

Genetic factors associated withACPA- RA

The number of genes that have been found to be significantly associated with ACPA- RA is much lower than for ACPA+ RA. This is probably due to the fact that only about 30 % of RA are ACPA- and only one GWAS analysis on a rather small material (774 patients and 1079 controls) has addressed specifically associations with ACPA- RA(71). Twin studies have shown that the heritability of the two forms of RA is similar, about 66% (72). Many genes associated with ACPA- RA remainthus to be

defined. It is characteristic for the polymorphism of these genes that they often are associated also with other autoimmune diseases. The heritability contributed by HLA is relatively weak. A recent study in UK showed a weak association with HLA-DR-SE alleles in RF positive, ACPA negative RA, but no association with double negative RA (73). In Japanese, one study suggests several HLA-DR associations in ACPA- RA (74) another study found no HLA association in ACPA- RA (44). Below we will describe some of the non-HLA genes that have been identified in ACPA- RA.

Neuropeptide S receptor 1 (NPSR1)(7p14.3)

The protein is expressed on various epithelial cells and is a member of the G protein-coupled receptor 1 family. Increased expression in ciliated cells of the respiratory epithelium and in bronchial smooth muscle cells is associated with asthma. Polymorphisms of the gene are associated with inflammatory bowel disease. The SNP rs324987 is associated with ACPA- RA and another SNP (rs10263447) seems to be associated with disease activity (75).

Interferon regulatory factor 5 (IRF5) (7q32)

IRF5 has already been described above, as it is also associated with ACPA+ RA at least for some SNPs (68). Several SNPs were associated preferentially with ACPA- RA, and IRF5 polymorphisms are also known to be associated with SLE.

CLEC16A (16p13.1)

This gene is expressed in immune cells and codes for a C-type lectin with unknown function. Polymorphisms of the gene are associated with other autoimmune diseases such as type-1 diabetes, multiple sclerosis, juvenile arthritis and Addison's disease. The SNP rs6498196 is associated with ACPA- RA and not with ACPA+ RA in a Norwegian population (76).

Combined impact of susceptibility genes, homozygosity effect and gene interactions

Although more than 35 genetic loci for RA susceptibility have now been defined, the combined heritability of the loci amounts to less than 20 % (39). Apart from the HLA system, the odds ratios for association with other loci are small and only become significant in large materials. Some of the missing heritability may be explained by homozygosity of recessive susceptibility alleles (77). Another factor to take into account is possible interactions between genes. Some examples of this phenomenon are given below. A more systematic analysis of multiple gene-gene interactions for RA candidate genes have been presented as part of genetic analysis workshop 16 (78).

Interaction PTPN22 (1p13)/HLA-DRB1

The *PTPN22* gene has already been described. A significant interaction with *HLA-DRB1-SE+* was found in three large case-control studies (79).

Interaction glutathione s-transferase Mu1 deletion (GSTM1-null, 1p13.1)/HLA-DRB1:

The glutathione s-transferase participates in the elimination of reactive oxygen species and other toxins. About 50 % of Europeans have a homozygous deletion of the *GSTM1* gene. A significant additive interaction of *GSTM1*-null and *HLA-DRB1-SE+* alleles for ACPA positivity has been shown in two cohorts(80).

Interaction between 5-hydroxytryptamine (serotonin) receptor 2A (HTR2A)(13q14-q21)/HLADRB1

Serotonin is a neurotransmitter, which however also is important for regulation of immune responses. Activation of serotonin receptors on antigen presenting cells reduces the production of TNF- α . A significant additive interaction between *HTR2A* TC haplotype and *HLA-DRB1-SE+* alleles for ACPA+ RA was observed in 3 independent cohorts and the results were confirmed by a metaanalysis(81).

Genes and environment

A multitude of environmental factors influence RA development. Smoking is strongly associated with ACPA+ RA, as shown in many studies and summed up in a metaanalysis (82). This may be explained by an increased citrullination of proteins in the upper airways provoked by the smoke. Another factor is periodontitis, because the pathogen *Porphyromonas gingivalis* has the capacity to citrullinate host peptides by secretion of peptidylarginine deiminase (83). Other indications on possible environmental factors may be found in a recent review on RA pathogenesis (22).

A word of caution in interpretation of genetic associations¹

Genetic analyses generally assume that a multifactorial disease such as RA is due to numerous genetic factors with minor, independent and cumulative effects (the disease appears above a certain threshold), and to environmental influences. Calculations involving a whole range of SNPs are done under this hypothesis. This model, the so-called polygenic model, along with the concept of

¹The European Society for Human Genetics has adopted a text concerning multifactorial diseases and this section is an extract of the text:
(<https://www.eshg.org/fileadmin/www.eshg.org/documents/received/2010MultifactorialDiseases.pdf>).

heritability was introduced by Fisher in 1918. In such a model, heritability of a trait measures the contribution of genetic variability to total variability of the trait, within a population. It is based on the hypothesis that genetic factors do not interact either with each other or with the environmental factors. However, for human diseases such as RA, this hypothesis is false or at least unrealistic as there obviously exist complex processes which underlie the pathophysiology, and environmental factors are still poorly described. Many GWAS interpretation neglect to recall such important aspects of the model applied. We have shown that, in addition, there is a huge gap between the detection of a significant OR and the identification of the genetic variations actually involved in the RA pathological process. Given the limited advances into understanding the mechanisms underlying the HLA associations observed over 35 years, testing hypotheses like the role of specific amino-acid positions for peptide binding as described above is important. In RA as in other multifactorial diseases, GWAS have revealed low-intensity signals (low odds ratio for the associated SNPs) thanks to impressive size of patient samples used. Such numbers imply also that little selection was carried out during sampling in terms of clinical and environmental homogeneity. As a consequence, it is very likely that the etiological heterogeneity is very high in these samples.

Attempting to measure "missing heritability" is actually misleading as there is no way of measuring either the importance or the nature of the missing information. A risk may be completely changed by just one new piece of environmental information. A strong heritability does not imply that environment only plays a minor role in the physiopathology of a disease. To illustrate this, an example can be taken in a different domain: on the basis of a genome analysis, someone could be declared to be at risk of developing leprosy although there is absolutely no risk if he or she is not exposed to the mycobacterium. According to those who promote the concept of missing heritability, it would reveal the "ground remaining to be covered" for the identification of all the genetic risk factors. However this might be a wrong approach and the exploration of gene-gene and gene-environment interactions using well characterized cohorts of patients in terms of clinical and environmental factors might be the route to go rather than multiplying large-scale GWAS.

This discussion recalls that, although GWAS and other studies are key elements in identifying new genes involved and in suggesting new relevant pathways to explore in order to understand RA, the estimated risks themselves are of poor value at individual level. This is especially important to underline as a number of companies propose genetic tests directly to consumers via internet, many of the tests being based on GWAS results and using risks calculations for establishing individual susceptibilities; several of them offer tests for susceptibility to RA (84). In the light of the above discussion of the interpretation and significance of risks derived from SNP testing caution is necessary towards such tests as their clinical utility is not established.

Conclusions

A large number of genes influences the development of RA. Most of the defined genes are associated with ACPA+ RA that seems to be a disease entity different from ACPA- RA. The association with HLA-DR is strong for ACPA+ RA, where specific residues characterize HLA-DR antigens that are associated with RA, probably because citrullinated peptides fit well into the antigen-binding groove. In disposed individuals, T and B cell autoimmunity to citrullinated peptides and proteins develops and this is probably an important pathogenic factor. As shown in the figure 3, most of the genes associated with ACPA+ RA are involved in T cell activation or in pathways in immune cells. Fewer genes have been identified for ACPA- RA susceptibility, and the genes are often at the same time involved in other autoimmune diseases. Further studies of the genetics of RA and the functional importance of susceptibility genes are important because they may give clues to how this invalidating autoimmune disease can be treated optimally.

Acknowledgements

AC is supported by an interface contract between Inserm and the CHU Toulouse and ARW is supported by the Association Midi-Pyrénées Santé, project APOGEE.

REFERENCES

1. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum.* 1987;30:1205-13.
2. Raychaudhuri S, Sandor C, Stahl EA, Freudenberg J, Lee HS, Jia X, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. *Nat Genet.* 2012;44:291-6.
3. Consortium WTCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature.* 2007;447:661-78.
4. Szodoray P, Szabo Z, Kapitany A, Gyetvai A, Lakos G, Szanto S, et al. Anti-citrullinated protein/peptide autoantibodies in association with genetic and environmental factors as indicators of disease outcome in rheumatoid arthritis. *Autoimmun Rev.* 2010;9:140-3.
5. Silman AJ, Pearson JE. Epidemiology and genetics of rheumatoid arthritis. *Arthritis Res.* 2002;4 Suppl 3:S265-72.
6. Scott DL, Wolfe F, Huizinga TW. Rheumatoid arthritis. *Lancet.* 2010;376:1094-108.

7. Sebbag M, Chapuy-Regaud S, Auger I, Petit-Teixeira E, Clavel C, Nogueira L, et al. Clinical and pathophysiological significance of the autoimmune response to citrullinated proteins in rheumatoid arthritis. *Joint Bone Spine*. 2004;71:493-502.
8. 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.
9. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, 3rd, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum*. 2010;62:2569-81.
10. Constantin A. Régulation génétique de la production des auto-anticorps au cours de la polyarthrite rhumatoïde. *Revue du rhumatisme monographies*. 2010;77:293-99.
11. van de Stadt LA, de Koning MH, van de Stadt RJ, Wolbink G, Dijkmans BA, Hamann D, et al. Development of the anti-citrullinated protein antibody repertoire prior to the onset of rheumatoid arthritis. *Arthritis Rheum*. 2011;63:3226-33.
12. 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.
13. 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.
14. Nishimura K, Sugiyama D, Kogata Y, Tsuji G, Nakazawa T, Kawano S, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. *Ann Intern Med*. 2007;146:797-808.
15. Farragher TM, Lunt M, Plant D, Bunn DK, Barton A, Symmons DP. Benefit of early treatment in inflammatory polyarthritis patients with anti-cyclic citrullinated peptide antibodies versus those without antibodies. *Arthritis Care Res (Hoboken)*. 2010;62:664-75.
16. Vossenaar ER, Zendman AJ, van Venrooij WJ, Pruijn GJ. PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *BioEssays : news and reviews in molecular, cellular and developmental biology*. 2003;25:1106-18.
17. Makrygiannakis D, Hermansson M, Ulfgren AK, Nicholas AP, Zendman AJ, Eklund A, et al. Smoking increases peptidylarginine deiminase 2 enzyme expression in human lungs and increases citrullination in BAL cells. *Annals of the rheumatic diseases*. 2008;67:1488-92.
18. van Beers JJ, Willemze A, Stammen-Vogelzangs J, Drijfhout JW, Toes RE, Pruijn GJ. Anti-citrullinated fibronectin antibodies in rheumatoid arthritis are associated with HLA-DRB1 shared epitope alleles. *Arthritis research & therapy*. 2012;14:R35.

19. Ireland JM, Unanue ER. Autophagy in antigen-presenting cells results in presentation of citrullinated peptides to CD4 T cells. *J Exp Med*. 2011;208:2625-32.
20. Feitsma AL, van der Voort EI, Franken KL, el Bannoudi H, Elferink BG, Drijfhout JW, et al. Identification of citrullinated vimentin peptides as T cell epitopes in HLA-DR4-positive patients with rheumatoid arthritis. *Arthritis Rheum*. 2010;62:117-25.
21. Bellatin MF, Han M, Fallena M, Fan L, Xia D, Olsen N, et al. Production of autoantibodies against citrullinated antigens/peptides by human B cells. *J Immunol*. 2012;188:3542-50.
22. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *The New England journal of medicine*. 2011;365:2205-19.
23. Cantaert T, Brouard S, Thurlings RM, Pallier A, Salinas GF, Braud C, et al. Alterations of the synovial T cell repertoire in anti-citrullinated protein antibody-positive rheumatoid arthritis. *Arthritis Rheum*. 2009;60:1944-56.
24. Steinman L. Mixed results with modulation of TH-17 cells in human autoimmune diseases. *Nature immunology*. 2010;11:41-4.
25. Genovese MC, Van den Bosch F, Roberson SA, Bojin S, Biagini IM, Ryan P, et al. LY2439821, a humanized anti-interleukin-17 monoclonal antibody, in the treatment of patients with rheumatoid arthritis: A phase I randomized, double-blind, placebo-controlled, proof-of-concept study. *Arthritis Rheum*. 2010;62:929-39.
26. Marston B, Palanichamy A, Anolik JH. B cells in the pathogenesis and treatment of rheumatoid arthritis. *Current opinion in rheumatology*. 2010;22:307-15.
27. Edwards JC, Szczepanski L, Szechinski J, Filipowicz-Sosnowska A, Emery P, Close DR, et al. Efficacy of B-cell-targeted therapy with rituximab in patients with rheumatoid arthritis. *The New England journal of medicine*. 2004;350:2572-81.
28. Firestein GS. Evolving concepts of rheumatoid arthritis. *Nature*. 2003;423:356-61.
29. Lu MC, Lai NS, Yu HC, Huang HB, Hsieh SC, Yu CL. Anti-citrullinated protein antibodies bind surface-expressed citrullinated Grp78 on monocyte/macrophages and stimulate tumor necrosis factor alpha production. *Arthritis Rheum*. 2010;62:1213-23.
30. Cascao R, Rosario HS, Souto-Carneiro MM, Fonseca JE. Neutrophils in rheumatoid arthritis: More than simple final effectors. *Autoimmun Rev*. 2010;9:531-5.
31. Hueber AJ, Asquith DL, Miller AM, Reilly J, Kerr S, Leipe J, et al. Mast cells express IL-17A in rheumatoid arthritis synovium. *J Immunol*. 2010;184:3336-40.
32. McInnes IB, Schett G. Cytokines in the pathogenesis of rheumatoid arthritis. *Nature reviews Immunology*. 2007;7:429-42.
33. Pesu M, Laurence A, Kishore N, Zwillich SH, Chan G, O'Shea JJ. Therapeutic targeting of Janus kinases. *Immunological reviews*. 2008;223:132-42.

34. Kremer JM, Bloom BJ, Breedveld FC, Coombs JH, Fletcher MP, Gruben D, et al. The safety and efficacy of a JAK inhibitor in patients with active rheumatoid arthritis: Results of a double-blind, placebo-controlled phase IIa trial of three dosage levels of CP-690,550 versus placebo. *Arthritis Rheum.* 2009;60:1895-905.
35. Bradfield PF, Amft N, Vernon-Wilson E, Exley AE, Parsonage G, Rainger GE, et al. Rheumatoid fibroblast-like synoviocytes overexpress the chemokine stromal cell-derived factor 1 (CXCL12), which supports distinct patterns and rates of CD4+ and CD8+ T cell migration within synovial tissue. *Arthritis Rheum.* 2003;48:2472-82.
36. Filer A, Parsonage G, Smith E, Osborne C, Thomas AM, Curnow SJ, et al. Differential survival of leukocyte subsets mediated by synovial, bone marrow, and skin fibroblasts: site-specific versus activation-dependent survival of T cells and neutrophils. *Arthritis Rheum.* 2006;54:2096-108.
37. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, Thomson BP, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. *Nat Genet.* 2010;42:508-14.
38. Raychaudhuri S, Plenge RM, Rossin EJ, Ng AC, Purcell SM, Sklar P, et al. Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS genetics.* 2009;5:e1000534.
39. Kurreeman FA, Stahl EA, Okada Y, Liao K, Diogo D, Raychaudhuri S, et al. Use of a Multiethnic Approach to Identify Rheumatoid- Arthritis-Susceptibility Loci, 1p36 and 17q12. *American journal of human genetics.* 2012;90:524-32.
40. Stastny P, Fink CW. HLA-Dw4 in adult and juvenile rheumatoid arthritis. *Transplantation proceedings.* 1977;9:1863-6.
41. Thomsen M, Morling N, Snorrason E, Svejgaard A, Sorensen SF. HLA--Dw4 and rheumatoid arthritis. *Tissue antigens.* 1979;13:56-60.
42. Freed BM, Schuyler RP, Aubrey MT. Association of the HLA-DRB1 epitope LA(67, 74) with rheumatoid arthritis and citrullinated vimentin binding. *Arthritis Rheum.* 2011;63:3733-9.
43. Mitsunaga S, Suzuki Y, Kuwana M, Sato S, Kaneko Y, Homma Y, et al. Associations between six classical HLA loci and rheumatoid arthritis: a comprehensive analysis. *Tissue antigens.* 2012.
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. *American journal of human genetics.* 2004;75:330-7.
45. Rantapaa-Dahlqvist S. What happens before the onset of rheumatoid arthritis? *Current opinion in rheumatology.* 2009;21:272-8.

46. Freudenberg J, Lee HS, Han BG, Shin HD, Kang YM, Sung YK, et al. Genome-wide association study of rheumatoid arthritis in Koreans: population-specific loci as well as overlap with European susceptibility loci. *Arthritis Rheum.* 2011;63:884-93.
47. Huang JJ, Qiu YR, Li HX, Sun DH, Yang J, Yang CL. A PTPN22 promoter polymorphism -1123G>C is associated with RA pathogenesis in Chinese. *Rheumatology international.* 2012;32:767-71.
48. Kurreeman F, Liao K, Chibnik L, Hickey B, Stahl E, Gainer V, et al. Genetic basis of autoantibody positive and negative rheumatoid arthritis risk in a multi-ethnic cohort derived from electronic health records. *American journal of human genetics.* 2011;88:57-69.
49. Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, Guiducci C, et al. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. *Nat Genet.* 2009;41:1313-8.
50. Suzuki A, Yamada R, Chang X, Tokuhira 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.
51. Burr ML, Naseem H, Hinks A, Eyre S, Gibbons LJ, Bowes J, et al. PADI4 genotype is not associated with rheumatoid arthritis in a large UK Caucasian population. *Annals of the rheumatic diseases.* 2010;69:666-70.
52. Chen JY, Wang CM, Ma CC, Hsu LA, Ho HH, Wu YJ, et al. A transmembrane polymorphism in FcγRIIb (FCGR2B) is associated with the production of anti-cyclic citrullinated peptide autoantibodies in Taiwanese RA. *Genes and immunity.* 2008;9:680-8.
53. Thabet MM, Huizinga TW, Marques RB, Stoeken-Rijsbergen G, Bakker AM, Kurreeman FA, et al. Contribution of Fcγ receptor IIIA gene 158V/F polymorphism and copy number variation to the risk of ACPA-positive rheumatoid arthritis. *Annals of the rheumatic diseases.* 2009;68:1775-80.
54. Robinson JI, Barrett JH, Taylor JC, Naven M, Corscadden D, Barton A, et al. Dissection of the FCGR3A association with RA: increased association in men and with autoantibody positive disease. *Annals of the rheumatic diseases.* 2010;69:1054-7.
55. Docampo E, Rabionet R, Riveira-Munoz E, Escaramis G, Julia A, Marsal S, et al. Deletion of the late cornified envelope genes, LCE3C and LCE3B, is associated with rheumatoid arthritis. *Arthritis Rheum.* 2010;62:1246-51.
56. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, Kastner DL, et al. REL, encoding a member of the NF-κappaB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nat Genet.* 2009;41:820-3.

57. Barton A, Eyre S, Ke X, Hinks A, Bowes J, Flynn E, 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. *Human molecular genetics*. 2009;18:2518-22.
58. 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. *The New England journal of medicine*. 2007;357:977-86.
59. Daha NA, Kurreeman FA, Marques RB, Stoeken-Rijsbergen G, Verduijn W, Huizinga TW, et al. Confirmation of STAT4, IL2/IL21, and CTLA4 polymorphisms in rheumatoid arthritis. *Arthritis Rheum*. 2009;60:1255-60.
60. 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. *American journal of human genetics*. 2007;81:1284-8.
61. Malysheva O, Pierer M, Wagner U, Wahle M, Baerwald CG. Association between beta2 adrenergic receptor polymorphisms and rheumatoid arthritis in conjunction with human leukocyte antigen (HLA)-DRB1 shared epitope. *Annals of the rheumatic diseases*. 2008;67:1759-64.
62. Pawlik A, Kurzawski M, Szczepanik T, Dziedziejko V, Safranow K, Borowiec-Chlopek Z, et al. Association of allograft inflammatory factor-1 gene polymorphism with rheumatoid arthritis. *Tissue antigens*. 2008;72:171-5.
63. Pawlik A, Kurzawski M, Dziedziejko V, Safranow K, Paczkowska E, Maslinski W, et al. Allograft Inflammatory Factor-1 Gene Polymorphisms in Patients with Rheumatoid Arthritis. *Genetic testing and molecular biomarkers*. 2011.
64. Scherer HU, van der Linden MP, Kurreeman FA, Stoeken-Rijsbergen G, Cessie S, Huizinga TW, et al. Association of the 6q23 region with the rate of joint destruction in rheumatoid arthritis. *Annals of the rheumatic diseases*. 2010;69:567-70.
65. Patsopoulos NA, Ioannidis JP. Susceptibility variants for rheumatoid arthritis in the TRAF1-C5 and 6q23 loci: a meta-analysis. *Annals of the rheumatic diseases*. 2010;69:561-6.
66. 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.
67. Sigurdsson S, Padyukov L, Kurreeman FA, Liljedahl U, Wiman AC, Alfredsson L, et al. Association of a haplotype in the promoter region of the interferon regulatory factor 5 gene with rheumatoid arthritis. *Arthritis Rheum*. 2007;56:2202-10.
68. 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;58:1264-74.

69. Orozco G, Eyre S, Hinks A, Ke X, Wilson AG, Bax DE, et al. Association of CD40 with rheumatoid arthritis confirmed in a large UK case-control study. *Annals of the rheumatic diseases.* 2010;69:813-6.

70. 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.

71. Padyukov L, Seielstad M, Ong RT, Ding B, Ronnelid J, Seddighzadeh M, et al. A genome-wide association study suggests contrasting associations in ACPA-positive versus ACPA-negative rheumatoid arthritis. *Annals of the rheumatic diseases.* 2011;70:259-65.

72. van der Woude D, Houwing-Duistermaat JJ, Toes RE, Huizinga TW, Thomson W, Worthington J, et al. Quantitative heritability of anti-citrullinated protein antibody-positive and anti-citrullinated protein antibody-negative rheumatoid arthritis. *Arthritis Rheum.* 2009;60:916-23.

73. Mackie SL, Taylor JC, Martin SG, Wordsworth P, Steer S, Wilson AG, et al. A spectrum of susceptibility to rheumatoid arthritis within HLA-DRB1: stratification by autoantibody status in a large UK population. *Genes and immunity.* 2012;13:120-8.

74. Terao C, Ohmura K, Kochi Y, Ikari K, Maruya E, Katayama M, et al. A large-scale association study identified multiple HLA-DRB1 alleles associated with ACPA-negative rheumatoid arthritis in Japanese subjects. *Annals of the rheumatic diseases.* 2011;70:2134-9.

75. D'Amato M, Zucchelli M, Seddighzadeh M, Anedda F, Lindblad S, Kere J, et al. Analysis of neuropeptide S receptor gene (NPSR1) polymorphism in rheumatoid arthritis. *PLoS one.* 2010;5:e9315.

76. Skinningsrud B, Lie BA, Husebye ES, Kvien TK, Forre O, Flato B, et al. A CLEC16A variant confers risk for juvenile idiopathic arthritis and anti-cyclic citrullinated peptide antibody negative rheumatoid arthritis. *Annals of the rheumatic diseases.* 2010;69:1471-4.

77. Yang HC, Chang LC, Liang YJ, Lin CH, Wang PL. A genome-wide homozygosity association study identifies runs of homozygosity associated with rheumatoid arthritis in the human major histocompatibility complex. *PLoS one.* 2012;7:e34840.

78. Huang CH, Cong L, Xie J, Qiao B, Lo SH, Zheng T. Rheumatoid arthritis-associated gene-gene interaction network for rheumatoid arthritis candidate genes. *BMC proceedings.* 2009;3 Suppl 7:S75.

79. Kallberg H, Padyukov L, Plenge RM, Ronnelid J, Gregersen PK, van der Helm-van Mil AH, et al. Gene-gene and gene-environment interactions involving HLA-DRB1, PTPN22, and smoking in two subsets of rheumatoid arthritis. *American journal of human genetics.* 2007;80:867-75.

80. Mikuls TR, Gould KA, Bynote KK, Yu F, Levan TD, Thiele GM, et al. Anticitrullinated protein antibody (ACPA) in rheumatoid arthritis: influence of an interaction between HLA-DRB1 shared epitope and a deletion polymorphism in glutathione S-transferase in a cross-sectional study. *Arthritis research & therapy*. 2010;12:R213.
81. Seddighzadeh M, Korotkova M, Kallberg H, Ding B, Daha N, Kurreeman FA, et al. Evidence for interaction between 5-hydroxytryptamine (serotonin) receptor 2A and MHC type II molecules in the development of rheumatoid arthritis. *European journal of human genetics : EJHG*. 2010;18:821-6.
82. Sugiyama D, Nishimura K, Tamaki K, Tsuji G, Nakazawa T, Morinobu A, et al. Impact of smoking as a risk factor for developing rheumatoid arthritis: a meta-analysis of observational studies. *Annals of the rheumatic diseases*. 2010;69:70-81.
83. Wegner N, Wait R, Sroka A, Eick S, Nguyen KA, Lundberg K, et al. Peptidylarginine deiminase from *Porphyromonas gingivalis* citrullinates human fibrinogen and alpha-enolase: implications for autoimmunity in rheumatoid arthritis. *Arthritis Rheum*. 2010;62:2662-72.
84. Ducournau P, Gourraud PA, Rial-Sebbag E, Bulle A, Cambon-Thomsen A. [Direct-to-consumer genetic testing through Internet: marketing, ethical and social issues]. *Medecine sciences : M/S*. 2011;27:95-102.
85. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*. 2004;25:1605-12.

Legends to figures

Fig 1

Post-translational conversion of the charged arginine residue to the uncharged citrulline residue.

The conversion is facilitated by the enzyme peptidyl arginine deiminase which is present in many cells and tissues in the body. The conversion is enhanced in lung tissues of smokers (17).

Fig 2

Ribbon model of HLA-DR.

The model has been elaborated based on the 3L6F entry in the Protein Data Bank. The peptide binding groove is shown and the four residues on the DR β chain that were associated with ACPA+RA in a large GWAS are highlighted (2). Another study in a mixed population suggested that also residue 67 is important, and this residue has been shown to interact with the binding of citrullinated vimentin peptide(42). The USCF Chimera program was used to create the figure (85).

Fig 3

Schematic representation of initiation of the auto-immune response in RA and the inflammatory phase.

The citrullination of peptides and proteins is an essential step in the initiation of the autoimmune response. A large number of the genes associated with RA code for surface molecules important for T-cell and B cell activation, as for example HLA, CD22, CD28, CD40, CD45, CD58, CTLA-4. Other genes code for cytokines or for factors involved in various pathways, as indicated in the boxes. See also table 1 and the text. After the initiation of an immune response towards citrullinated peptides and proteins, a clinical disease usually develops, with inflammatory responses towards the cells of the joints. Activated macrophages, neutrophils and cytokines are important actors in the destructive process. RANKL (receptor activator of NF- κ B ligand) has not been found to display a polymorphism associated with RA, but is important for the destruction of bone tissue via activation of osteoclasts.

Table 1: Genes and proteins associated with anti-citrullinated peptide antibody (ACPA) positive RA

<i>Gene (localization)</i>	Protein and its function	SNP or allele	Odds Ratio	95 % conf. interval	Ref(s)
<i>HLA-DRB1 (6p21.3)</i>	The polymorphic HLA-DRβ1 chain presents together with the non-polymorphic α chain antigenic peptides to CD4+ T cells. Amino acids at position 11, 13, 71 and 74 determine the susceptibility to ACPA+ RA.	HLA-DRB1*04:01	4.44	4.02-4.91	(2)
		HLA-DRB1*04:04, *04:05, *04:08, *10:01	4.22	3.75-4-75	(2)
		HLA-DRB1*01:01, *01:02	2.17	1.94-2.42	(2)
		HLA-DRB1*16:01	2.04	1.59-2.62	(2)
		HLA-DRB1*04:03, *04:07	1.65	1.24-2.19	(2)
		HLA-DRB1*09:01	1.65	1.29-2.10	(2)
		HLA-DRB1*04:02	1.43	1.04-1.96	(2)
<i>HLA-B (6p21.3)</i>	HLA class I molecules present antigenic peptides to CD8+ T cells	HLA-B*08	2.12	1.89-2.38	(2)
<i>HLA-DPB1 (6p21.3)</i>	The polymorphic HLA-DPβ1 chain presents together with the non-polymorphic α chain antigenic peptides to CD4+ T cells	Several alleles with Phe in position 9	1.40	1.31-1.50	(2)
<i>PTPN22 (1p13)</i>	Protein-tyrosine phosphatase (LYP) is a lymphocyte-specific intracellular phosphatase involved in regulation of lymphocyte activation	rs2476601	1.94	1.81-2.08	(37)
		rs6679677	2.06	1.68-2.53	(48)
<i>CD2, CD58 (1p13)</i>	CD2 is a surface antigen expressed on T cells and NK cells CD58 is the ligand of CD2	rs11586238	1.13	1.07-1.19	(37)
<i>TNFRSF14 (1p36)</i>	TNFRSF14 is a member of TNF receptor superfamily mediating signals induced by TNF receptor associated factors (TRAF)	rs3890745	1.12	1.06-1.18	(37)
<i>PADI4 (1p36.13)</i>	PAD is isoform 4 of a family of 5 enzymes responsible for the conversion of arginine residues to citrulline residues	Association in Asians but not in Europeans, see text			(46, 50, 51)
<i>FCGR (1q21-23)</i>	FcγR are receptor proteins for the Fc part of IgG, involved in the process of phagocytosis and clearing of immune complexes	FCGR2B rs1050501	1.35 (Taiwanese)	1.09-1.66	(52)
		FCGR3A rs393991	1.3	1.01-1.55	(53)
		FCGR2A rs12746613	1.13	1.06-1.21	(49)

<i>LCE3C_LCE3B-del</i> (1q21.3)	Late cornified envelope genes located in the epidermal differentiation complex	Homozygous deletion of gene associated with rs4112788	1.51	1.09-2.13	(55)
<i>PTPRC</i> (1q31.3)	PTPRC is member of the protein tyrosine phosphatase (PTP) family, CD45	rs10919563	1.14	1.06-1.22	(37)
<i>REL</i> (2p13)	c-Rel is member of NF-κB family important for Th1 immune responses	rs13031237	1.13	1.07-1.18	(37)
<i>SPRED2</i> (2p14)	SPRED2 regulates activation of CD45+ hematopoietic cells via the Ras-MAP kinase pathway	rs934734	1.13	1.06-1.21	(37)
<i>AFF3</i> (2q11)	AFF3 is a tissue restricted nuclear transcriptional activator preferentially expressed in lymphoid tissue.	rs11676922	1.12	1.07-1.17	(37)
<i>STAT4</i> (2q32)	STAT4 is member of the STAT family of transcription factors	rs7574865	1.16	1.10-1.23	(37)
<i>CD28</i> (2q33)	CD28 is a surface protein essential for T-cell proliferation and survival	rs1980422	1.12	1.06-1.18	(37)
<i>CTLA4</i> (2q33)	CTLA-4 is a surface protein on T cells which transmits an inhibitory signal	rs3087243	1.15	1.10-1.20	(37)
<i>PXK</i> (3p14)	PXK modulates Na,K-ATPase enzymatic and ion pump activities	rs13315591	1.13	1.04-1.23	(37)
<i>RBPJ</i> (4p15)	RBPJ is a transcriptional regulator important in the Notch signaling pathway	rs874040	1.18	1.12-1.24	(37)
<i>IL2, IL21</i> (4q27)	IL-2 and IL21 are cytokines important for cellular immune responses	rs6822844	1.11	1.05-1.19	(37)
<i>ANKRD55, IL6ST</i> (5q11)	The function of ANKRD55 is unknown IL6ST as a signal transducer shared by cytokines as IL-6	rs10040327 rs6859219	1.28 1.18	1.05-1.57 1.08-1.28	(44) (37)
<i>ADRB2</i> (5q32-q34)	ADBR2 is a Beta-2-adrenergic receptor	Arg16	4.43	2.81-7.02	(61)
<i>AIF-1</i>	AIF-1 is a cytoplasmic inflammation responsive protein expressed in joints from RA	rs2269475	8.8	2.1-37.7	(62)

(6p21.3)	patients	rs2259571	2.30	1.25-4.25	(63)
<i>PRDM1</i> (6q21)	PRDM1 acts as a repressor of beta-interferon gene expression	rs548234	1.10	1.05-1.16	(37)
<i>TNFAIP3</i> (6q23)	The TNFAIP3 protein inhibits NF-kappa B activation as well as TNF-mediated apoptosis	rs10499194	1.10	1.04-1.15	(37)
		rs6920220	1.22	1.16-1.29	(37)
		rs5029937	1.40	1.24-1.58	(37)
<i>TAGAP</i> (6q25)	TAGAP has a function as a Rho GTPase-activating protein	rs394581	1.10	1.04-1.15	(37)
<i>CCR6</i> (6q27)	CCR6 is member of the beta chemokine receptor family expressed by immature dendritic cells and memory T cells	rs3093023	1.11	1.06-1.16	(37)
<i>IRF5</i> (7q32)	IRF5 is member of the interferon regulatory factor (IRF) family, a group of transcription factors with diverse roles	rs10488631	1.25	1.14-1.37	(37)
<i>BLK</i> (8p23)	BLK is nonreceptor tyrosine-kinase of the src family of proto-oncogenes	rs13277113	1.15	1.00-1.33	(44)
		rs2736340	1.12	1.07-1.18	(37)
<i>CCL21</i> (9p13)	CCL21 is a chemokine that inhibits hemopoiesis and stimulates chemotaxis,	rs2812378	1.10	1.05-1.16	(37)
<i>TRAF1/C5</i> (9q33-34.1)	The TRAF1 protein associates with and mediates the signal transduction from various receptors of the TNFR superfamily C5 is the fifth component of complement	rs3761847	1.13	1.08-1.18	(37)
		rs2900180	1.28	1.13-1.44	(65)
		rs10118357	1.19	1.05-1.34	(48)
<i>IL2RA</i> (10p15)	The interleukin 2 receptor alpha chain is part of the IL-2 receptor	rs706778	1.11	1.06-1.17	(37)
<i>PRKCC</i> (10p15)	Protein kinase C theta is one of the PKC family members	rs4750316	1.15	1.09-1.22	(37)
<i>RAG1, TRAF6</i> (11p12)	RAG1 is Involved in activation of immunoglobulin V-D-J recombination TRAF6 is member of the TNF receptor associated factor protein family	rs540386	1.16	p=6.1E-04	(49)
<i>IKZF3</i> (17q12)	(IKAROS family zinc finger 3) participates in regulation and proliferation of B cells	rs2872507	1.10	p=1.7E-09	(39)

<i>CD40</i> (20q13)	CD40 is member of the TNF-receptor superfamily, essential in mediating a broad variety of immune and inflammatory responses	rs4810485	1.23	1.11-1.37	(69)
------------------------	---	-----------	------	-----------	------

The genes are sorted according to their chromosomal localization, except HLA genes that are on the top. Most of the data are derived from genome wide association studies (GWAS) on large materials and concern patients and controls of European ancestry, except where indicated. As detailed in the text, most of the genes are involved in T-cell activation, either coding for surface molecules on immune cells, cytokines or proteins participating in signal transduction in immune cells.