NON-HLA SUSCEPTIBILITY GENES AND POLYMORPHISMS IN HUNGARIAN RHEUMATOID ARTHRITIS PATIENTS

Ph.D. thesis

BERNADETT FARAGÓ

Supervisor: Béla Melegh M.D., Ph.D., D.Sc.

Department of Medical Genetics and Child Development, Faculty of Medicine,
University of Pécs

PÉCS, 2011.

1. INTRODUCTION

1.1. Autoimmune diseases in general

Autoimmunity is a biological phenomenon, an essential part of the mechanism of life. Despite the fact that autoimmune reactions are continuously occurring in an organism – also at molecular, cell and physiogical levels – sometimes severe or even fatal diseases based on abnormal immune-regulation (mostly due to the dysfunction of autoreactive T cells and autoantibodies) may occur.

Autoimmunity works as a result of a network-like regulation, depending on several components. The defect of any component included, or the gene it is coded by, can lead to the alteration of the normal function of the immune system, and to the development of an autoimmune disease.

All autoimmune diseases are characterized by the abnormal immune-tolerance against own antigens, and all of them are associated with certain genetic factors.

Autoimmune diseases can be divided into two groups: organ-specific diseases, like diabetes mellitus, intestinal bowel diseases (ulcerative colitis, Crohn's disease), psoriasis or autoimmune thyroiditis affect only one organ system, while systemic diseases, like systemic lupus erythematosus, systemic sclerosis, Sjögren's syndrome and rheumatoid arthritis are spread through the whole body.

1.2. Rheumatoid arthritis

1.2.1. The clinical manifestation of rheumatoid arthritis

Rheumatoid arthritis (RA) is the most frequent systemic autoimmune disease affecting approximately 1% of the human population including Hungary. RA is present in the whole world, but in Africa it occurs more rare. RA shows its highest prevalence among North-American inhabitants (Pima and Chippewa tribes). According to the observation RA affects approximately 2.5 times more women, than men.

Rheumatoid arthritis is manifested with the chronic inflammation of small joints, becoming gradually even more destructive. In severe forms RA might also have extraarticular symptoms. Without appropriate treatment, patients might be hindered in their normal lifestyle, even disability and permanent disablement can develop. Symptoms are usually present symmetrically on both sides of the body. Most frequent symptoms are pain, stiffness of the joints, swollen peripheral joints, regional osteoporosis and ankylosis, although the clinical manifestation might be quite heterogeneous.

Symptoms can arise at any age however they usually develop between the age of 40 and 70. Joint destruction of synovial origin usually come forward within the first two years of the disease and can be diagnosed using radiographic methods. MRI helps in the detection of synovial hypertrophies, bone oedema and other early erosive alterations already within the first four months.

1.2.2. Serological characteristics

The diagnosis of RA is based on both imaging and laboratory techniques, complemented by the evaluation of the clinical symptoms. Although several serological

tests are available, two factors of high sensitivity and specifity for RA should be discussed. Both are easily measured from the peripheral blood.

Rheumatoid factor (RF) is an autoantibody, directly binding to the Fc-region of normal human immunoglobuline-G (IgG) molecule, but its binding to other immunoglobuline-isotypes (IgE, IgM, IgA) has also been observed. In clinical diagnostics, the measurement of the IgM-RF immune-complex is most commonly applied. This complex is at present in 75% of RA patients, while the other 25% is negative for the test, despite they do show several specific symptoms of the disease. Additionally, RF can occur even in patients not suffering from RA, but having certain infections or other rheumatologic diseases. According to the observations, smoking and coffee consumption can facilitate the accumulation of RF. The detection of RF already before the onset of RA-specific clinical symptoms is possible only in 19.3% of all cases. Taking all these into account, the specifity and the prognostic value of RF is rather low. However, the severity and rapidity of destruction can be easily estimated in the light of RF-serology. Usually, RF-positive patients develop a more severe type of the disease and are affected even by extraarticular manifestations.

As a conclusion, RF-testing using ELISA should not be regarded as a screening method, but can be useful when other symptoms alluding to RA are observed in a patient.

The relatively low specifity of RF has lead to the investigation of other, more specific and sensitive RA-autoantibodies.

Anti-cyclic citrullinated peptide (anti-CCP) autoantibodies are produced against peptides that become citrullinated in the course of the activity of specific enzymes. Anti-CCP autoantibodies are accumulated directly within the inflamed region: they are located in the synovium. Anti-CCP shows a slightly lower sensitivity than RF (66.4%), but it is characterized with a much higher specifity (97.1%) than that. This high sensitivity is caused by the RA-specific mechanism of the dysregulation of humoral immune response against citrullinated peptides. Anti-CCP can be detected even in early RA (in 40-60% of the cases), and is also present in 34.5% of RF-negative patients, particularly in the early phase of RA. The presence of anti-CCP might allude to the later onset of severe joint destruction and progressive procession of the disease.

The coexistence of anti-CCP and RF predestinates the development of RA with the probability of 28.9 times, compared to the healthy population negative for both serofactors.

1.2.3. The etiology of rheumatoid arthritis

Similarly to other autoimmune diseases, the coexistence of several different factors is responsible for the precipitation of RA. Besides genetic predisposition, an infection caused by a microbial organism and also real autoimmune mechanisms play important roles in the generation of the disorder. As RA is usually accompanied with a chronic inflammation, it should be regarded as a so called „immunoinflammatoric" disease, rather than a real autoimmune disorder.

1.2.3.1. Environmental factors

Several studies on the possible environmental factors contributing to the etiology of RA have been published, although the findings are rather conflicting: while smoking seems to facilitate the manifestation of RA, alcohol is suspected to have an opposite effect. The predisposition to RA is also raised by coffee consumption, oral contraceptives, obesity and schizophrenia among first-degree relatives. However the relevance of these factors is highly dependent on the serological status of the patients (e.g.: anti-CCP positivity). According to a study performed in Denmark, only the age of

women at the time of menarche shows a significant correlation with the development of RA, independent from any serological characteristics.

A bacterial or viral infection behind RA is also commonly suspected: mostly Mycobacteria, Sptreptococcus species, Proteus and Chlamydia species, Parvovirus B19, Rubella and HIV viruses have been investigated, although none of them proved to undoubtedly effect the evolving of RA. It is suspected, that these infectious agents predominate only in patients carrying certain genetic conditions.

1.2.3.2. Genetic factors

Monozygotic and dizygotic twin studies suggest that the significance of the genetic factors in the etiology of RA is approximately 60%. The environmental factors discussed above might be responsible for the other 40%.

The presence of certain variants of the human leukocyte antigen gene complex (HLA) explains approximately half of the genetic predisposition, while several other and independent genes and their polymorphisms contribute to the formation of the overall genetic predisposition. Sometimes the coexistence of more variants is needed to develop the predisposing genetic conditions.

1.2.3.2.1. The *HLA* gene-complex

While investigating the genetic background of rheumatoid arthritis and other autoimmune diseases, the *HLA*-gene complex was the first to be proven to influence the development and the process of RA.

The MHC (major histocompatibility complex) molecules coded by the *HLA*-genes can be divided into three groups. Regarding RA, the molecules of group II (MHC-II) should be discussed. These are cell-surface proteins with immunoglobulin structure, and play role in antigen-presentation. Their role is essential in initiating both humoral and cellular immune responses. MHC-II molecules are mainly expressed in macrophages, B cells and activated T cells.

The gene loci coding MHC-II molecules are nominated as *HLA-DR, DQ* and *DP*, and each locus have several alleles. *HLA*-genes show a strong linkage to each other, therefore they are usually inherited as a unit (haplotype).

The MHC-molecules associated with RA have a certain and much conservated sequence consisting of 5 amino acids in the third hipervariablility region of the DRB-chain. This sequence is called the shared epitope, and is suspected to be carried by microorganisms playing role in the development of RA. During antigen-presentation these epitopes are suspected to be presented with a particular affinity.

Latest investigations has however revealed, that the shared epitope hypothesis by Gregersen discussed above, does not undoubtedly stand for RA, as 45% of the Caucasian population do carry the shared epitope without developing RA ever.

1.2.3.2.2. Other susceptibility genes

The heterogenic clinical manifestations of rheumatoid arthritis suggested that besides HLA several other genetic factors and their polymorphisms might play a role in developing the genetic predisposition of RA.

Among the identified genetic agents, mainly genes coding molecules of particular significance (cytokines, kemokines and their receptors) can be found. Some of these polymorphisms are located in the following genes: *interleukin-1 (IL-1), IL-1 promoter, tumor-nekrosis-factor-α (TNF α) promoter, interferon-γ (IFNγ), IL-3, IL-4, IL-6, IL-10, IL-12,* a *CCR5* kemokine-receptor, *p53 tumor-supressor*, az *NFKBIL1* (Nuclear factor kappa-

B inhibitor-like 1 gene), and *STAT4* (signal transducer and activator of transcription 4 gene).

2. AIMS

2.1. Genetic variants

The aim of our work was to investigate the frequency and the possible predisposing nature of genes and polymorphisms already suspected to correlate to RA, but which had not been investigated in Hungarian populations earlier. We tried to select genes the role of which in the pathogenesis of RA seemed not to be obvious: their significance might depend on the population examined or are only relevant in certain subpopulation of RA patients.

2.1.1. Haplotypes encoded by PADI4 gene variants

The peptidylarginine deiminase citrullinating enzymes (PADIs; E.C. 3.5.3.15.) are involved in the post-translational deimination of arginine residues to citrulline in proteins. Citrullination partially unfolds the proteins via loss of the positive charge in the arginine moiety, which can affect the antigenicity of the protein chains. PADIs play a specific role in the pathogenesis of rheumatoid arthritis, as citrullinated proteins are the targets of anti-citrullinated peptide antibodies (ACPAs), including that against the cyclic citrullinated peptide (anti-CCP), which is the most sensitive RA-specific autoantibody. The antibodies are usually generated at an early stage of the disease.

Amongst the naturally occurring variants of the *PADI4* gene, some have been found to confer susceptibility to RA in Asian populations. Recent studies on Europeans, including British, French and Spanish Caucasian populations, could not confirm these findings. Only one German case-control study showed an association of a functional haplotype with the disease.

Variations in the amino acid sequence of the PADIs can influence their immunological features; these variants can have different immune responses and ultimately the coded characters can thereby affect the production of ACPAs. The aim of this present study was to: (i) defi ne the haplotypes and their frequencies in a Hungarian population of RA patients; (ii) test if any of the haplotypes can confer susceptibility for RA in the average population; and (iii) study the haplotype distribution in serologically characterized subgroups and to test thereby the possible associations of the haplogroups with anti-CCP or RF positivity, alone or in combination.

2.1.2. *PTPN22* gene, 1858C/T variant

It is well-known that different autoimmune disorders may share common susceptibility loci. One of these loci is a protein tyrosine phosphatase gene (PTPN22; *MIM 600716*), which is expressed in lymphocytes and encodes the intracellular protein tyrosine phosphatase nonreceptor type 22, also known as lymphoid-specific phosphatase (Lyp). The C1858T allele of this gene has been suggested to increase the risk for developing several autoimmune diseases, like type 1 diabetes, systemic lupus erythematosus, Grave's disease, Hashimoto thyroiditis and rheumatoid arthritis. The C1858T substitution results in an arginine to tryptophan change in codon 620 of the protein product. This non-synonymous change affects the proline-rich SH3 binding site

of the PTPN22 protein, and can modify its binding to the intracellular c-src protein tyrosine kinase (Csk). Normally, these molecules act together in inactivating the lymphocyte-specific protein tyrosine kinase, involved in the negative regulation of the early T cell activating systems. Thus, the *PTPN22* 1858T variant can be responsible for the pathogenic immune-response in inflammatory diseases like RA, via the affected T cell activation process.

Controversial results have been published about the significance of rheumatoid factor and of anti-CCP seropositivity in relation with the role of C1858T variant in the development of RA. The aim of our work was to characterize these relationships in serologically stratified Hungarian RA patients.

2.1.3. *CTLA4* gene, +49A/G andCT60 variants

The cytotoxic T lymphocyte associated antigen (CTLA4) is a well-known co-signalling molecule of the B7 receptor immunoglobulin family. CTLA4 is expressed by CD4+, CD8+, T and B lymphocytes, and is located on the surface of the cell. Its expression and its density on the cell surface is affected by the activation of the T lymphocytes. Contrary to CD28, which transmits a stimulatory signal to the T cells by binding to B7-1 or B7-2 receptors on antigen presenting cells, CTLA4 is an inhibitory factor in T-cell activating systems.

The *CTLA4* gene polymorphisms are reportedly associated with a variety of T cell mediated autoimmune diseases such as type 1 diabetes mellitus, systemic lupus erythematosus, Hashimoto thyroiditis, Addison disease, Grave's disease, multiple sclerosis and celiac disease. Previous investigations on the association of *CTLA4* with RA resulted in controversial outcomes.

Nistico *et al.* identified a +49A/G transition polymorphism of the leader sequence in exon 1, resulting in a threonine to alanine change at position 17 of the amino acid sequence. In patients carrying the homozygous GG genotype the hydrophobic alanine is inserted into the highly conservative sequence of the CTLA4 leader peptide. This slight alteration of the leader peptide may affect the intracellular transport of the CTLA4 protein and its availability on the cell surface. As GG homozygotes have fewer functionally active CTLA4 molecules on the cell surface of the T lymphocytes, fewer B7-CTLA4 associations can form, leading to a less efficient CTLA4-mediated T cell inhibition. As a consequence, a significantly increased T cell proliferation can be observed, compared to AA homozygotes.

Another polymorphism, the CT60 (A6230G) allele in the 3'-UTR region encodes either a protective A variant or a predisposing G variant for autoimmune diseases. The G allele is associated with lower mRNA levels of the soluble CTLA4 isoform. Both the +49A/G and the CT60A/G variants have been reported to confer risk for rheumatoid arthritis, however, the results are controversial, some findings affirm the role of the SNPs in RA, while others could not reveal any association. The aim of our study was to examine the distribution of these two SNPs in Hungarian RA patients separately and in haplotype combinations.

2.1.4. *IL23R* gene, rs10889677C/A, rs2201841T/C and rs1884444G/T variants

Interleukin-23 is a heterodimeric proinflammatory cytokine composed of a p19 subunit and the p40 subunit, which one is also part of IL-12. IL-23 plays a central regulatory role in the differentiation of native CD4+ T cells into IL-17-producing T helper cells. IL-23 also acts on neutrophils and macrophages to induce the generation of other inflammatory cytokines, such as IL-1, IL-6, and TNF-α. Furthermore, it has an effect on CD8+ T cells and γδ T cells that also produce IL-17. Taken together, IL-23 plays a unique

function in the initiation and perpetuation of innate and T cell-mediated inflammation. IL-23 affects memory T cells and inflammatory macrophage function through the engagement of the IL23R.

The *IL23R* gene located on chromosome 1p31 encodes one subunit of the receptor. In a recent study, Duerr et al. found an association between this *IL-23* receptor gene and inflammatory bowel diseases, and they reported several independent functional single nucleotide polymorphisms (SNPs) of the gene and its neighbouring region which either conferred strong protection against, or marked susceptibility to Crohn's disease and ulcerative colitis in non-Jewish subjects. In previous studies using genetically deficient animals, experimental evidence has been presented that IL-23 has also unique features in the autoimmune destruction-associated events not only in inflammatory bowel diseases, but also in collagen-induced arthritis. Therefore, we hypothesized that functional variants of the *IL23R* gene can have also role in human RA, and we studied the prevalence of selected representatives of the *IL23R* gene in this disease group. From the *IL23R* variants reported by Duerr et al. for Crohn's disease we chose the rs10889677 allele, which affects the exon-3'UTR of the gene, and the rs2201841, which is an SNP in intron 7; both of them have been reported as susceptibility genes. Besides them, the distribution of the neutral rs1884444 allele, which associates with His3Gln change in the extracellular domain, was also determined.

3. MATERILAS AND METHODS

3.1. Patients

The DNA samples of the patient group were obtained from a central pool governed by our Department as part of the Central National Biobank Network of Hungary (www.biobank.hu). We have collected the blood samples of 452 rheumatoid arthritis patients, but we do hope that this number will be constantly increasing. Approximately half of the samples (n=236) was collected by the University of Pécs, while 216 other samples were provided by the University os Debrecen. In general, the DNA samples were derived from the average Caucasian Hungarian population; the minorities were thereby randomly included at their natural distribution rate. Only patients fulfilling the diagnostic criteria of the American College of Rheumatology for rheumathoid arthritis were selected for the study.

The control group consisted of 297 clinically healthy individuals, who had no evidence of any clinical, laboratory or history records for any metabolic disorder or systemic illness, including autoimmune diseases. They all were blood donor volunteers, students and staff members of our departments. The age and gender distribution of the control group was matched to the similar characteristics of the patient group.

3.2. Serology testing

The sera from non-hemolyzed blood of each RA-patient and control was tested for the presence of RF- and anti-CCP autoantibodies using the Rheumatoid Factor Screen ORG522S test by ORGENTEC Diagnostika GmbH (Mainz, Germany) and the enzyme immunoassay of Euro-Diagnostica (Malmö, Sweden), respectively. Only subjects negative for the presence both RF- and anti-CCP were allowed to serve as controls.

3.3. Genotype analysis

Genomic DNA was extracted from peripheral blood leukocytes using a routine desalting method. The DNA analyses started with the amplification of the target

sequences, applying polymerase chain reaction (PCR) method. Amplification was performed in MJ Research PTC 200 thermal cyclers under appropriate conditions, in the presence of synthetic sequence specific oligonucleotide primers, Taq-polymerase, dNTPs, buffer and genomic DNA-template. The amplified products were separated by agarose gel electrophoresis and visualized by ethidium bromide staining. The variants of the genes in our focus (*PADI4, PTPN22, CTLA4, IL23R*) were genotyped performing specific restriction fragment length polymorphism (RFLP) methods. During the entire study period the guidelines and regulations approved by the local Research-Ethical Committee of the Medical and Healthyscience College of Pecs in 10.07.2000., 04.02.2003., and 09.03.2004. were followed.

3.3.1. Haplotypes encoded by PADI4 gene variants

Genomic DNA was extracted from peripheral blood leukocytes using a routine desalting method. We examined four exonic PADI4 SNPs: padi4_89 (163G/A, GenBank rs11203366), and padi4_90 (245T/C, GenBank rs11203367) in exon 2, padi4_92 (335C/G, GenBank rs874881) in exon 3 and padi4_104 (349C/T, GenBank rs1748033) in exon 4 and two intronic SNPs of the same gene, padi4_94 (17535226C/T on chromosome 1, GenBank rs2240340) and padi4_102 (17546809C/T on chromosome 1, GenBank rs2240337). The nomenclature followed is that of Suzuki *et al*. The naturally occurring haplotypes are listed in Table 1. Each of the exonic genetic variants are associated with an amino acid change in the protein products: a Gly55Ser, Val82Ala, Gly112Ala, and Leu117Leu modification, respectively.

The following primers were designed and used to amplify the examined sequences: for padi4_89*G/A forward 5'-CTC CTC ACT GCA TCC TCT GCT-3', reverse 5'-CTT TCA TCG TCA GGG TCA CCT CTA-3'; for padi4_90*T/C forward 5'-CAA AGT CCC ACG ATC TGC AAG-3', reverse 5'-AGG ACA CTA TGG CTG GAA GAA GC-3'; for padi4_92*G/C forward 5'-AGC TTT TTG CTT TCC CTC CAT T-3' and reverse 5'-GTC TGA CTG GCT AGA AAC CAT GC-3'; for padi4_94*C/T forward 5'-CTC ACC AAC CTC TCC TGG TAC-3' and reverse 5'-TCA CCA ATT GTG GGT TCA GA-3'; for padi4_102*C/T forward 5'-CTG GCC CAG GCA CCA CCA G-3' and reverse 5'-AGG GTT TCG GCA GCT GTG CC-3', and for padi4_104*C/T forward 5'-CAT CAC AGT TGT GGC CCC G-3' and reverse 5'-GCG GGT GAT GTC TGC GCC C-3'. The mismatch bases are underlined. The PCR amplifi cations were performed on MJ Research PTC 200 thermal cyclers according to the following protocol: initial denaturation at 95ºC for 2 min followed by 35 cycles of denaturation at 94ºC for 30 sec, annealing at 55ºC for 30 sec (at 60ºC in the case of padi4_102 and padi4_104, and at 57ºC in the case of padi4_94) and extension at 72ºC for 30 sec. Final extension was carried out at 72ºC for 5 min. The amplicons were digested with allele-specific restriction endonucleases, *HaeIII* for padi4_89, *MlsI* for padi4_90 *HpaII* for padi4_92, *KpnI* for padi4_94, *RsaI* for padi4_102 and *PasI* for padi4_104. In all amplicons there was an obligatory cleaving site to enable us to control the efficacy of the digestion.

3.3.2. *PTPN22* gene, 1858C/T variant

The C1858T functional variant (*GenBank rs2476601*) was examined using a PCR-RFLP-assay. The specific primer pair used for the amplification of the target sequence was the following: 5'-TTT TAG ACA TCA AAT GTT GCT CAG-3' (forward) and 5'-AAG AGA ATT TAT TTT GCT TTT TCC-3' (reverse). The PCR amplification was performed on MJ Research PTC200 thermal cyclers using the appropriate reaction conditions. The 688 bp amplicons were digested with *Rsa I* (or with the isoschisomer *Afa I*) restriction endonuclease. The amplicon contained an obligatory cleavage site to control the efficacy of the enzyme digestion. For allele C fragments with 44, 223 and 421 bp lenghts

could be detected; while for the T allele 267 and 421 bp fragments were detectable. In heterozygous subjects (TC) all four different fragments were present.

3.3.3. *CTLA4* gene, +49A/G and CT60 variants

The +49A/G (*GenBank rs231775*) and the CT60 A/G alleles (*GenBank rs3087243*) were examined by PCR RFLP methods. For the amplification of the target sequences the following primers were designed: 5'-CTT GAG GTT GTC TTT TCG AG-3' (forward) and 5'-TAC TAA ATA CCT GGC GCT CT-3' (reverse) primers for +49G/A; and 5'-ATC TGT GGT GGT CGT TTT CC-3' (forward) and 5'-TGG AAA CCA AAT GTG CTG AG-3' (reverse) primers for CT60. The PCR amplifications were performed on MJ Research PTC 200 thermal cyclers using the following conditions: initial denaturation at 96ºC for 3 min followed by 35 cycles of denaturation at 96ºC for 45 sec, annealing at 56ºC for 45 sec and extension at 72ºC for 45 sec. The final extension was 10 minutes at 72ºC.

The amplicons were then digested by restriction endonucleases, +49A/G by *Bse XI* and CT60 by *Mae II*, as recommended by the manufacturer (Fermentas International Inc., Burlington, Ontario, Canada). In both methods designed the amplicon contained an obligate restriction cleavage site allowing us to monitor the efficacy of the enyzme digestion. After agarose gel electrophoresis and visualization under UV light the following fragments could be detected: 39 bp and 541 bp for the +49A/G A allele; 39 bp, 261 bp and 274 bp fragments for the G allele; 202 bp and 403 bp for the CT60 A allele, and 151, 202 and 252 bp long fragments in the case of the G allele. In heterozygotes, all the possible fragments of the examined variants were at present.

3.3.4. *IL23R* gene, rs10889677C/A, rs2201841T/C and rs1884444G/T variants

PCR-RFLP methods were applied to test the alleles of the *IL-23* receptor gene (GenBank NM_144701, GeneID 149233), using the following forward and reverse primers: 5'-ATC GTG AAT GAG GAG TTG CC-3' and 5'-TGT GCC TGT ATG TGT GAC CA-3' for rs10889677; 5'-GGC AAA AGG GAA TTG AGA GG-3' and 5'-GGC CTA TGA TTA TGC TTT TTC CTG-3' for rs2201841; and 5'-CAG TCT TTT CCT GCT TCC AGA CAT GAA TC-3' and 5'-AAT AAA ATC ATA CTC TTG CCA ATG GCC C-3' for rs1884444 alleles.

For RFLP tests *MnlI, HpyF3I* and *PscI* restriction endonucleases were used, respectively; each primer sets were designed to make an obligate cleavage site on the amplicon to enable us to control the efficacy of the digestion. In two primers we had to introduce mismatch bases to generate artificial cleaving sites (underlined in the sequence). The digestion of the amplicon of rs10889677C resulted in 61, 185 and 225 bp bands, while the A allele was indicated by 185 and 286 bp digestion products. The rs2201841T allele resulted in 163 and 257 bp fragments, while the C allele in 25, 163 and 232 bp bands. The rs1884444G allele was characterized by bands with 191 and 318 bp, while for the T allele 28, 191 and 290 bp digestion products could be detected.

3.4. Statistical analysis

Statistical analysis was carried out using SPSS 11.5 for Windows. We performed binary logistic regression analyses and chi-square tests to reveal the possible associations between the genetic and serological characteristics.

During the linkage studies of the examined *IL23R* gene polymorphisms the Haploview 4.1. program was used. Results are expressed with the R^2 correlation coefficient and the D' variate.

4. RESULTS

4.1. Haplotypes encoded by PADI4 gene variants

The allele frequencies for all the *PADI4* single nucleotide variants of the six examined loci were in Hardy-Weinberg equilibrium both in the RA and the control subjects. None of the variants showed significant accumulation in RA patients compared to controls (Table 2). For padi4_102, the prevalence rate of the T allele was signifi cantly higher in the control group than in the patients with RA (15.2% *vs* 24.4%; $\chi 2 = 5.05$; $p = 0.025$; OR = 0.54; 95%CI: 0.32-0.93). The haplotype frequencies are shown in Table 2. Haplotype 1 (characterized by padi4_89*A, padi4_90*C, padi_92*C, padi4_94*C, padi4_102*C and padi4_104*C) and haplotype 2 (padi4_89*G, padi4_90*T, padi4_92*G, padi4_94*T, padi4_102*C and padi4_104*T) were the most frequent. We also detected haplotype 1B (classification by Hoppe *et al*), which is formed from three different haplotypes determined by the six SNPs studied. Coexistence of the padi4_89*A, padi4_90*C and padi4_92*G alleles was common in these. Frequency of the 1B haplotype did not differ significantly between cases and controls. There was no accumulation of any haplotypes in the RA patients compared with the controls. None of the haplotypes showed an increased frequency in the RF- and anti-CCP-positive RA subjects (Table 3). Moreover, no difference was observed in the distribution of the *PADI4*-haplotypes in patients with combined seropositivity (RF plus anti-CCP), nor in patients with combined seronegativity.

4.2. *PTPN22* gene, 1858C/T variant

Table 4 shows the distribution of the PTPN22 C1858T genotypes and the T allele observed in patients and controls. The alleles were in Hardy-Weinberg equilibrium in all examined groups.

The frequency of the TC genotype and the carriage of the 1858T variant, defined as the carriage of TC+TT genotypes together was significantly increased in the overall RA population and the specific seropositive subgroups (characterized by patients with the presence of RF- and/or anti-CCP). A similar increase in the prevalence of the TT genotype was observed in these groups. Besides, we could not detect a statistically significant accumulation of the T variant in the subpopulation seronegative for both examined serofactors.

Regression analyses using an unadjusted model, and models adjusted to age and gender were performed to evaluate the possible correlation between the 1858T allele and the disease susceptibility. Carriage of the T allele was associated with the development of RA using the crude assay (data not shown); and the correction for age and gender also confirmed the T allele as independent susceptibility factor (Table 4). A marked gene dosage effect was verified for the 1858T genotype. The binary regression analysis used to study the whole RA population and the serologically different subsets of patients revealed, that the association of the TT genotype with the development of RA was more than twice as strong as found for the heterozygous T genotype (Table 4).

4.3. *CTLA4* gene, +49A/G and CT60 variants

The study comprised of 428 cases with RA and 230 healthy controls. Upon isolation of genomic DNA from all patients and controls, we performed genotyping using PCR-RFLP methods specific for *CTLA4* variants +49 A/G and CT60.

Table 5 presents our findings, the allele frequencies of the CT60 and +49A/G SNPs and of the +49*G-CT60*G haplotype in the patients and controls. The CT60*GG genotype and the CT60*G allele frequencies showed a significant increase in the RF- and/or anti-CCP-seropositive RA-populations, while no difference was observed in the distribution of the +49A/G alleles compared to the healthy controls. All alleles studied here were in Hardy-Weinberg equilibrium in both the patient and control groups.

We also tested the frequencies of the possible haplotypic combinations and their significance. Comparing the prevalences of the +49*A – CT60*A , +49*A – CT60*G, +49*G – CT60*A , and +49*G – CT60*G combinations, the +49*G - CT60*G haplotype was found to accumulate in the entire RA patient population, and further analysis revealed that it was present at significantly elevated frequencies in the seropositive groups (Table 5).

The bivariate logistic regression analyses adjusted to age and gender revealed that the CT60*G and CT60*GG variants represent independent risk for the development of RF-, anti-CCP-seropositive and combined types of RA (Table 6). Similarly, the simultaneous presence of the CT60*G and +49*G alleles showed an approximately 1.5-fold increase in disease susceptibility considering the RF- and/or anti-CCP-seropositive cases. The +49*G variant alone was not found to confer risk for the disease.

4.4.*IL23R* gene, rs10889677C/A, rs2201841T/C and rs1884444G/T variants

The genotyping of the three *IL23R* variants under our scope was performed on 412 rheumatoid arthritis patients (96 male and 316 female, average age of 54,9 ± 14,7 years) and 220 controls (115 male and 105 female, average age of 41,7 ± 13,1 years). The prevalence rates of the examined *IL23R* variants are shown in Table 7. All the genotype and allele distributions were in Hardy-Weinberg-equlibrium both in the patient groups and in the controls.

The rs10889677 AA genotype was increased by approximately 2-fold in the RA population; similar prevalence rates were observed in subjects with RF, anti-CCP, and combined RF plus anti-CCP seropositivity. The logistic regression analysis revealed, that the AA genotype represents an independent risk factor for the development of rheumatoid arthritis.

Similarly, for the rs2201841 allele the prevalence of the CC genotype showed more than 2-fold increase in the RA groups compared to the controls. RF, anti-CCP, and combined RF plus anti-CCP seropositivity did not affect allele distribution. Regression analysis model revealed, that similarly to the rs10889677 variant, carrying the rs2201841 CC genotype alone confers a more than 2-fold risk for RA.

The AA rs10889677 and CC rs2201841 alleles were simultaneously present in 9.95% of RA patients, while this rate was only 3.64% in the controls (p=0.008), data not shown).

Contrary to the rs10889677 and rs2201841 alleles, there was no difference in the distribution of any variants of rs1884444 allele in any group of patients.

Table 1: The length of restriction fragments derived from the digestion of the examined PADI4 variants (base-pairs)

PADI4 SNP	ALLELE 1		ALLELE 2	
padi4_89	G	40 115	A	40 155
padi4_90	T	108 177 593 770	C	108 770
padi4_92	C	100 285	G	100 134 151
padi4_94	C	21 87 204	T	108 204
padi4_102	C	40 359	T	36 40 323
padi4_104	C	22 98 169	T	120 169

Table 2: Haplotype structure and frequency in the PADI4 gene.

Haplotype ID	SNP ID (padi4_x)						Haplotype frequency				Regression analysis			
	89	90	92	94	102	104	Case (n = 334)		Control (n = 260)		χ^2	p	OR	(95%CI)
Haplotype 1	A	C	C	C	C	C	170	(50.9)	125	(48.1)	0.47	0.495	1.12	(0.81-1.55)
Haplotype 2	G	T	G	T	C	T	92	(27.5)	60	(23.1)	1.53	0.216	1.27	(0.87-1.84)
Haplotype 3	G	T	G	T	T	T	20	(5.98)	16	(6.15)	0.01	0.933	0.97	(0.49-1.91)
Haplotype 4	G	T	G	T	C	C	28	(8.38)	24	(9.23)	0.13	0.717	0.90	(0.51-1.59)
Other	NA	NA	NA	NA	NA	NA	29	(8.68)	34	(13.1)	NA	NA		NA

Values represent the numbers of cases, with the relative frequencies (expressed as a percentage) between parentheses; in the last column the ranges of odds ratios are given between parentheses. Chi-squares and ORs have been calculated for the haplotype frequencies in cases vs. controls. NA: not applicable.

Table 3: Haplotype distribution and frequencies of PADI4 SNPs in RA patients positive for rheumatoid factor (RF), anti-CCP or both.

	RA patients					
	RF		Anti-CCP		RF and anti-CCP	
	Positive (n = 242)	Negative (n = 92)	Positive (n = 214)	Negative (n = 120)	Positive-positive (n = 200)	Negative-negative (n = 78)
Haplotype 1	125 (51.7)	45 (48.9)	108 (50.5)	62 (51.7)	102 (51.0)	39 (50.0)
Haplotype 2	68 (28.1)	24 (26.1)	63 (29.4)	29 (24.2)	59 (29.5)	20 (25.6)
Haplotype 3	13 (5.37)	7 (7.61)	12 (5.61)	8 (6.67)	10 (5.00)	5 (6.41)
Haplotype 4	22 (9.09)	6 (6.52)	19 (8.88)	9 (7.50)	18 (9.00)	5 (6.41)
Other	14 (5.79)	10 (10.9)	12 (5.61)	12 (10.0)	11 (5.50)	9 (11.5)

Values represent the numbers of cases, with the relative frequencies (expressed as a percentage) between parentheses.

Table 4: *Prevalence of the PTPN22 C1858T allele variants in patients with RA and in the controls.*

	RA-PATIENTS				CONTROL
	Total n=399	RF-positive n=299	anti-CCP-positive n=286	RF + anti-CCP-positive n=263	Total n=107
CC (%)	241 (60.4)	178 (59.5)	171 (59.8)	161 (61.2)	85 (79.4)
TC (%)	121 (30.3)	92 (30.8)	86 (30.1)	76 (28.9)	20 (18.7)
TC+TT (%)	158 (39.6)*	121 (40.5)*	115 (40.2)*	102 (38.8)*	22 (20.6)
TT (%)	37 (9.27)*	29 (9.69)*	29 (10.1)*	26 (9.89)*	2 (1.87)
T (%)	195 (24.4)*	150 (25.1)*	144 (25.2)*	128 (24.3)*	24 (11.2)
OR (95% CI) for TC+TT	2.48	2.57	2.55	2.39	
vs. controls	(1.48-4.18)	(1.51-4.38)	(1.49-4.36)	(1.39-4.11)	
p	0.001	0.001	0.001	0.002	
OR (95% CI) for TC vs.	1.89	2.00	1.88	1.81	
controls	(1.10-3.24)	(1.15-3.49)	(1.07-3.28)	(1.03-3.19)	
p	0.022	0.015	0.027	0.040	
OR (95% CI) for TT vs.	5.04	4.71	5.49	4.94	
controls	(1.18-21.5)	(1.09-20.3)	(1.27-23.7)	(1.14-21.4)	
p	0.029	0.038	0.022	0.033	

* p at least <0.05 vs. controls

*Table 5: The distribution and frequencies of the CTLA4 CT60 and +49A/G genotypes and alleles, and the +49*G-CT60*G haplotype stratified by RF- and anti-CCP-seropositivity in patients versus controls.*

	RA patients							Control
		RF		anti-CCP		RF + anti-CCP		
	Total ($n = 428$)	Positive ($n = 324$)	Negative ($n = 104$)	Positive ($n = 310$)	Negative ($n = 118$)	Positive-positive ($n = 286$)	Negative-negative ($n = 80$)	Total ($n = 230$)
CT60								
AA (%)	84 (19.6)*	62 (19.1)*	22 (21.2)	59 (19.0)*	25 (21.2)	54 (18.9)*	17 (21.3)	64 (24.0)
GA (%)	199 (46.5)	143 (44.1)	56 (53.8)	131 (42.3)	68 (57.6)	122 (42.7)	47 (58.8)	109 (48.8)
GA+GG (%)	344 (80.4)	262 (80.9)	82 (78.8)	251 (81.0)	93 (78.8)	232 (81.1)	63 (78.8)	166 (78.4)
GG (%)	145 (33.9)*	119 (36.7)*	26 (25.0)	120 (38.7)*	25 (21.2)	110 (38.5)*	16 (20.0)	57 (27.1)
G (%)	489 (57.1)*	381 (58.8)*	108 (51.9)	371 (59.8)*	118 (50.0)	342 (59.8)*	79 (49.4)	223 (48.5)
+49A/G								
AA (%)	160 (37.4)	122 (37.7)	38 (36.5)	112 (36.1)	48 (40.7)	104 (36.4)	30 (37.5)	98 (42.6)
GA (%)	201 (47.0)	151 (46.6)	50 (48.1)	147 (47.4)	54 (45.8)	135 (47.2)	38 (47.5)	97 (42.2)
GA+GG (%)	268 (62.6)	202 (62.3)	66 (63.5)	198 (63.9)	70 (59.3)	182 (63.6)	50 (62.5)	132 (57.4)
GG (%)	67 (15.7)	51 (15.7)	16 (15.4)	51 (16.5)	16 (13.6)	47 (16.4)	12 (15.0)	35 (15.2)
G (%)	335 (39.1)	253 (39.0)	82 (39.4)	249 (40.2)	86 (36.4)	229 (40.0)	62 (38.8)	167 (36.3)
+49*G-CT60*G	182 (21.3)*	147 (22.7)*	35 (16.8)	147 (23.7)*	35 (14.8)	134 (23.4)*	22 (13.8)	62 (13.5)

*$P <$ .05 in the prevalence of the AA and GG genotypes, the G allele, or the +49*G-CT60*G haplotype versus controls.

*Table 6: The results of the binary logistic regression analyses considering the CT60*GG genotype, the CT60*G allele, and the +49*G-CT60 *G haplotype in RA patients stratified by RF- and anti-CCP-seropositivity versus controls.*

	RA Patients						
		RF		anti-CCP		RF + anti-CCP	
	Total	Positive	Negative	Positive	Negative	Positive-positive	Negative-negative
CT60*GG							
OR (95%CI)	1.56 (1.09–2.23)	1.76 (1.21–2.56)	1.01 (0.59–1.73)	1.92 (1.32–2.79)	0.82 (0.48–1.39)	1.90 (1.29–2.78)	0.76 (0.41–1.42)
P	.016	.003	.966	.001	.455	.001	.386
CT60*G							
OR (95%CI)	1.42 (1.13–1.78)	1.52 (1.20–1.93)	1.15 (0.83–1.60)	1.58 (1.24–2.02)	1.06 (0.78–1.46)	1.58 (1.23–2.03)	1.04 (0.72–1.49)
P	.003	.001	.410	< .001	.145	< .001	.845
+49*G - CT60*G							
OR (95%CI)	1.73 (1.27–2.37)	1.88 (1.36–2.61)	1.30 (0.83–2.04)	2.00 (1.44–2.76)	1.12 (0.71–1.75)	1.96 (1.41–2.73)	1.02 (0.61–1.73)
P	.001	<0 .001	.256	<0 .001	0.626	< .001	.931

Table 7: Interleukin-23 receptor exon-3'UTR rs10889677, intronic rs2201841 and exonic rs1884444 (His3Gln) genotypes in the groups of patients withthe studied diseases and in healthy controls (p<0.05 vs. controls using the χ2-test.)*

	RHEUMATOID ARTHRITIS				CROHN'S DISEASE	SYSTEMIC SCLEROSIS	CONTROLS
	Total	RF positive	Anti-CCP positive	RF- and anti-CCP positive			
rs10889677	n=412	n=310	n=293	n=272	n=190	n=224	n=220
CC	183 (44.4%)	135 (43.5%)	129 (44.0%)	118 (43.4%)	75 (39.5%)	114 (50.9%)	96 (43.6%)
AC	180 (43.7%)	136 (43.9%)	132 (45.1%)	123 (45.2%)	92 (48.4%)	102 (45.5%)	111 (50.5%)
AC+AA	229 (55.6%)	175 (56.5%)	164 (56.0%)	154 (56.6%)	115 (60.5%)	110 (49.1%)	124 (56.4%)
AA	49 (11.9%)*	39 (12.6%)*	32 (10.9%)*	31 (11.4%)*	23 (12.1%)*	8 (3.57%)	13 (5.91%)
OR (95% CI)	2.15	2.29	1.95	2.05	2.19	0.59	
for AA vs. CC+AC	(1.14-4.06)	(1.19-4.40)	(0.99-3.82)	(1.04-4.02)	(1.08-4.46)	(0.24-1.45)	
p	0.018	0.013	0.050	0.037	0.030	0.251	
rs2201841							
TT	178 (43.2%)	127 (41.0%)	122 (41.6%)	109 (40.1%)	75 (39.5%)	117 (52.2%)	101 (45.9%)
CT	180 (43.7%)	140 (45.2%)	133 (45.4%)	126 (46.3%)	90 (47.4%)	95 (42.4%)	106 (48.2%)
CT+TT	234 (56.8%)	183 (59.0%)	171 (58.4%)	163 (59.9%)	115 (60.5%)	107 (47.8%)	119 (54.1%)
CC	54 (13.1%)*	43 (13.9%)*	38 (13.0%)*	37 (13.6%)*	25 (13.2%)*	12 (5.36%)	13 (5.91%)
OR (95% CI)	2.40	2.56	2.37	2.51	2.41	0.90	
for CC vs. TT+CT	(1.28-4.51)	(1.34-4.89)	(1.23-4.57)	(1.30-4.85)	(1.20-4.86)	(0.40-2.02)	
p	0.006	0.004	0.010	0.006	0.014	0.801	
rs1884444							
GG	118 (28.6%)	78 (25.2%)	74 (25.3%)	65 (23.9%)	57 (30.0%)	54 (24.1%)	55 (25.0%)
GT	281 (68.2%)	220 (71.0%)	210 (71.7%)	198 (72.8%)	132 (69.5%)	163 (72.8%)	162 (73.6%)
GT+TT	294 (71.4%)	232 (74.8%)	219 (74.7%)	207 (76.1%)	133 (70.0%)	170 (75.2%)	165 (75.0%)
TT	13 (3.16%)	12 (3.87%)	9 (3.07%)	9 (3.31%)	1 (0.53%)	7 (3.13%)	3 (1.36%)
OR (95% CI)	2.36	2.91	2.29	2.48	0.38	2.33	
for TT vs. GG+GT	(0.66-8.36)	(0.81-10.4)	(0.61-8.57)	(0.66-9.26)	(0.04-3.71)	(0.60-9.14)	
p	0.185	0.101	0.218	0.178	0.407	0.224	

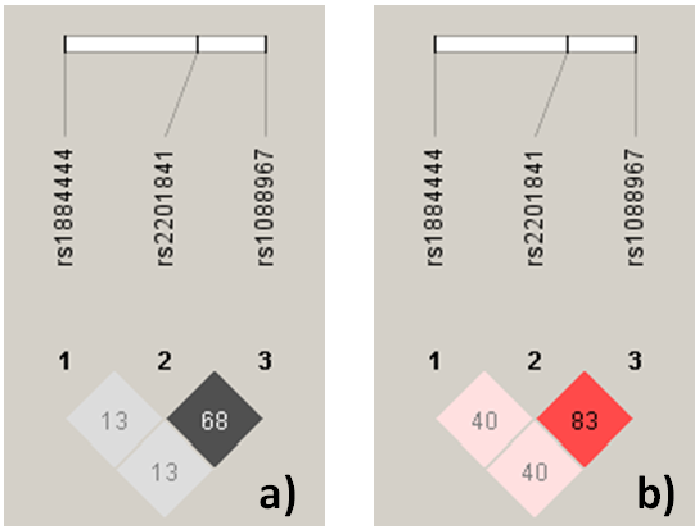

Figure 1: Linkage analysis of the three examined IL23R polymorphisms.
a) values of the R^2 correlation coefficient b) values of the D' variate

5. DISCUSSION

Rheumatoid arthritis is a rather common immunoinflammatoric disease, which is supposed to develop due to the coexistence of certain environmental, microbial and genetic factors. The genetic predisposition is responsible for the onset of RA in approximately 60%, half of which can be traced back to certain variants of the *HLA*-gene complex. The heterogeneous clinical manifestations of RA suggested, that several other genes might and undoubtedly do participate in the formation of genetic predisposition of the disease. We aimed to examine genes and polymorphisms in an average Hungarian RA-population and in matched controls, the role of which has still risen up several questions in the international fields of literature.

During the diagnostic work process we applied specific PCR/RFLP methods. Results were compared performing the chi-square test. We performed binary logistic regression analyses to reveal the possible associations between the genetic and serological characteristics.

Amongst the naturally occurring variants of the *PADI4* gene, some have been reported to confer susceptibility to RA in Asian populations, whereas examinations of European groups have failed to confirm these associations, except for a German cohort. One explanation of these differences could be the differing genetic structure of the populations examined. In this context, the Hungarians are unique in the Carpathian basin because a large proportion of the earliest inhabitants came from the east, beyond the Urals.

In the present study we did not confine ourselves to an examination of the distribution of PADI4 gene haplotypes in Hungarians, but also tested for possible correlations between haplotypes and seropositivity. Our findings on allele and haplotype frequencies are similar to those previously described in other European studies. This means that the Hungarian population under study here did not differ from other Europeans in this respect. No accumulation of any specific haplotype was observed in the patients with RA, which means that none of the major haplotypes confer susceptibility to the disease. Further analysis in patient groups characterized

for the presence of anti-CCP and/or RF antibodies also did not reveal the accumulation of any haplotypes in the seropositive RA patients.

In the current study the samples of unrelated RA-patients and matched healthy controls were genotyped for the C1858T variants of the PTPN22 gene. Similarly to previous studies on other Caucasian populations, the T variant showed an increased prevalence in the overall Hungarian group of RA-patients, compared to the healthy controls. The used regression analysis models revealed that the presence of the T allele conferred disease susceptibility in RA-patients.

A marked gene dosage effect was also observed in risk increase for the disease: the presence of a single T allele represented an approximately 2-fold risk, while in homozygotes with two copies of the T variant 4.9-5.5-fold ratios were observed. As it is well-known that some autoimmune diseases share common or similar molecular background, it is likely that the presumed molecular mechanisms behind the association of the 1858T variant and type 1 diabetes reported by Bottini et al also stand for RA: it is supposed that only the Lyp enzyme containing the 1858C (Arg620) variant is able to form a complex with Csk kinase, whereas the Lyp with 1858T (Trp620) is not. This results in individuals heterozygous for the 1858T allele in reduced amounts of Lyp-Csk complexes, while TT homozygotes have no such complexes at all. As the Lyp-Csk complex is responsible for the inhibition of T cell activation, the T cells lacking Lyp-Csk complexes become hyper-reactive, and are more likely to generate a destructive immune response against autoantigens after immune stresses.

In summary, the data presented here show, that in the Hungarians the C1858T variants of the PTPN22 gene represent susceptibility for the disease mainly in seropositive patients, although we cannot exclude its role in patients seronegative for both factors; and the association is gene dosage dependent.

We genotyped two functional SNPs of the *CTLA4* gene in RA-patients and matched controls, previously characterized for RF- and anti-CCP-seropositivity. The CT60*GG genotype and the CT60*G allele exhibited an accumulation in the entire RA-group when compared with the healthy controls (33.9% vs. 27.1% and 57.1% vs. 48.5%). However, subsequent analysis of the data revealed a clear serology dependency: all seropositive groups of patients showed a significantly increased prevalence of the CT60*GG genotype and the CT60*G allele compared to the controls. On the other hand, no similar differences could be established between the seronegative patients and the control cases regarding the CT60*GG genotype- and CT60*G allele-distribution. This suggests, that carrying the CT60*G variant does not mean predisposition for RF- and/or anti-CCP-seronegative subtypes of RA, however it might confer susceptibility for RF- and/or anti-CCP-seropositive forms of the disease.

In RF- and anti-CCP-seropositive subjects the CT60*G allele represented an approximately 1.5-fold risk for the development of RA in our patient population. These results are in agreement with previous findings. It has been reported that the CT60*G variant is likely associated with the RF-seropositive form of RA. So far no data has been published on samples seropositive both for RF and anti-CCP. Our examinations revealed a risk of 1.58 (95%CI: 1.23-2.03, $p<0.001$) in the double-seropositive group, which is similar to that observed in the subgroups characterized by seropositivity for either RF or anti-CCP (OR=1.52, 95%CI: 1.20-1.93, $p=0.001$; and OR=1.58, 95%CI: 1.24-2.02, $p<0.001$, respectively).

Based on the relatively high ratio of patients with the homozygous CT60*GG genotype we hypothesize that the CT*60G is a susceptibility factor for RA in the Hungarian population. Furthermore, the CT60*GG homozygous state represents an approximately 2-fold risk for RA in the seropositive subsets of samples, compared to the risk of carrying the G allele in a heterozygous form. This may suggest a gene dose dependency in influencing disease susceptibility, although the presence of two copies of the CT60*G variant does not mean an actual double risk, compared to a single copy of the variant.

Considering the +49A/G polymorphism, we could not detect any significant differences regarding the genotype and allele distributions of the RA and control samples. Previous works documented controversial outcomes on the possible association between this variant and RA. It is

possible, that the +49A/G polymorphism does not increase disease susceptibility alone, but together with other predisposing factors it increases the risk for developing RA, as some studies demonstrated an association in HLA-DRB1*04 patients, while others found no association irrespective of the HLA-shared epitope status. A study on Chinese patients suggested that CT60 is a possible independent causal variant in RA, and that the effect of the +49*G variant is dependent on the CT60*G allele, since when the +49*G allele is combined with the CT60*G in a haplotype, their effect is increased.

Based on recent findings by several research groups it is becoming increasingly clear that besides inflammatory bowel disease, IL-23 has also essential role in the autoimmune processes as well in the experimental rheumatoid arthritis. With this role of IL-23 in animal rheumatoid arthritis model in mind, we decided to study the prevalence of selected variants in an average human population with this disease. Similar to the Crohn's disease, the rs10889677AA and the rs2201841CC genotypes showed an accumulation in RA patients, the logistic regression analysis revealed susceptibility nature of these alleles in the development of the disease. The prevalence of the rs10889677AA and the rs2201841CC genotypes were similar in patients with RF, anti-CCP, or combined seropositivity, showing that the susceptibility-conferring character is not influenced by the presence or absence of the classic serology features. The relative high rate of accumulation of the AA rs10889677 and CC rs2201841 allelic variants in the same subjects suggest, that certain haplotypes might be responsible for the susceptibility nature of *IL23R*.

Although the exact explanation of the events remain to be determined for both Crohn's disease and rheumatoid arthritis, some possibilities come from the known functions of IL-23. IL-23 initiates and perpetuates both innate and T cell-mediated forms of certain tissue specific autoimmune disorders. If the IL-23/IL-17 axis becomes dysregulated, there is a risk for the development of autoimmune pathologies and autoimmune diseases. IL-23 also affects T cells and inflammatory neutrophil and macrophage function through action on the IL-23 receptor; all these elements can be possibly involved in the events leading to the development of these diseases.

Anti p40 therapies, which inhibit both IL-12 and IL-23, due to their shared p40 subunit, showed promising results in human trials for inflammatory bowel disease. With the recognition of the role of *IL-23* functional variants in inflammatory bowel disease, it is clear that further differentiations are required to verify the subgroup of patients who will benefit from the selective suppression of the IL-23 signaling. Similarly, with the current observations this can be extended also for rheumatoid arthritis as well, that would represent new perspective toward a personalized therapy of patients with this disease.

The results of the genetic examinations performed by our research group will hopefully contribute to the recognition and understanding of the genetic background of rheumatoid arthritis, which is crucial in the early diagnosis and in the improvement of new treatments for the disease. The identification of further predisposing genes and wide-range genotype-phenotype analyses are also highly required, as the improving knowledge can lead to effective prevention and treatment only by integrating the new aspects revealed using up-to-date methods like molecular genetic analysis.

6. CONCLUSION

i. Our results suggest that none of the examined *PADI4* gene polymorphisms (padi4_89, padi4_90, padi4_92, padi4_94, padi4_102, padi4_104) or their haplotypes mean predisposition to the disease in the Hungarian population of RA-patients.

ii. In the RF and/or anti-CCP seropositive types of rheumatoid arthritis, carrying either the *PTPN22* gene 1858T variant, or the *CTLA4* gene CT60*G allele means an independent predisposition to RA. 1858T allele predominates in a definite, while the CT60 G allele predominates in a moderate, gene-dose dependent manner.

iii. According to our findings, *CTLA4* gene +49A/G variant does not contribute to the genetic predisposition of RA alone, however together with the CT60*G allele it definitely raises the predisposing nature of the latter.

iv. Carrying any of the two variants out of the three examined polymorphisms (rs10889677, rs2201841, rs1884444) of the *IL23R* gene, the rs10889677 AA genotype and the rs2201841 CC genotype result in an excessive risk for developing RA. This predisposition is independent from the serological status (RF and/or anti-CCP-seropositivity) of the patient.

v. The frequent coexistence of the rs10889677 AA and rs2201841 CC genotypes and the linkage analyses performed suggest that the polymorphisms of the *IL23R* gene form haplotypes, and contribute to the genetic predisposition of rheumatoid arthritis together.

7. LIST OF PUBLICATIONS

7.1. The thesis is based on the following publications

1 **Faragó, B**., Talián, G.C., Maász, A., Magyari, L., Horvatovich, K., Kovács, B., Cserép, V., Kisfali, P., Kiss, C.G., Czirják, L. és Melegh, B.: Prevalence of functional haplotypes of the peptidylarginine deiminase citrullinating enzyme gene in patients with rheumatoid arthritis: no influence of the presence of anti-citrullinated peptide antibodies. Clin Exp Rheumatol. 2007(4):523-8. IF=2.270

2 **Farago, B**., Magyari, L., Safrany, E., Csongei, V., Jaromi, L., Horvatovich, K., Sipeky, C., Maasz, A., Radics, J., Gyetvai, A., Szekanecz, Z., Czirjak, L.és Melegh, B.: Functional variants of interleukin-23 receptor gene confer risk for rheumatoid arthritis but not for systemic sclerosis. Ann Rheum Dis. 2008;67(2):248-250. IF=7.188

3 **Farago, B**., Safrany, E., Melegh, B.: Role of interleukin-23 receptor polymorphisms in rheumatoid arthritis. In Press. Review. Hungarian.

4 **Farago, B**., Talian, G.C., Komlosi, K., Nagy, G., Berki, T., Gyetvai, A., Szekanecz, Z., Nyarady, Z., Kiss, C.G., Nemeth, P., Czirjak, L. és Melegh, B.: Protein tyrosine phosphatase gene C1858T allele confers risk for rheumatoid arthritis in Hungarian subjects. Rheumatol Int. 2009;29(7):793-796. IF=1.493

5 **Faragó, B**., Kisfali, P., Magyari, L., Polgár, N. és Melegh, B.: Cytotoxic T lymphocyte associated antigen +49G variant confers risk for anti-CCP- and rheumatoid factor-positive type of rheumatoid arthritis only in combination with CT60*G allele. Autoimmune Diseases. 2010. In Press.

7.2. Other publications

1 Szolnoki, Z., Maasz, A., Magyari, L., Horvatovich, K., **Farago, B**., Somogyvari, F., Kondacs, A., Szabo, M., Fodor, L., Bodor, A., Hadarits, F. és Melegh, B.: Coexistence of angiotensin II Type-1 receptor A1166C and angiotensin-converting enzyme D/D polymorphism suggests susceptibility for small-vessel-associated ischemic stroke. Neuromolecular Med. 2006;8:353-360. IF=3.396

2 Magyari, L., Bene, J., Komlosi, K., Talian, G., **Farago, B**., Csongei, V., Jaromi, L., Safrany, E., Sipeky, C., Lakner, L., Varga, M., Gasztonyi, B.és Melegh, B.: Prevalence of SLC22A4 1672T and SLC22A5 -207C Combination Defined TC Haplotype in Hungarian Ulcerative Colitis Patients. Pathol Oncol Res. 2007;13(1):53-6. IF=1.272

3 Magyari, L., **Faragó, B**., Bene, J., Horvatovich, K., Lakner, L., Varga, M., Figler, M., Gasztonyi, B., Mózsik, G. és Melegh, B.: No association of the cytotoxic T-lymphocyte associated gene CTLA4 +49A/G polymorphisms with Crohn's disease and ulcerative colitis in Hungarian population samples. World J Gastroenterol. 2007 Apr 21;13(15):2205-8.

4 Szolnoki, Z., Maasz, A., Magyari, L., Horvatovich, K., **Farago, B**., Somogyvari, F., Kondacs, A., Szabo, M., Bodor, A., Hadarits, F. és Melegh, B.: The combination of homozygous MTHFR 677T and angiotensin II type-1 receptor 1166C variants confers the risk of small-vessel-associated ischemic stroke. J Mol Neurosci. 2007;31(3):201-7. IF=1.735

5 Sáfrány, E., Balikó, L., Guseo, A., **Faragó, B**. és Melegh, B.: The autosomal dominant cerebellar ataxias are hereditary neurodegenerative diseases. Orv Hetil. 2007 Nov 11;148(45):2125-32. Review. Hungarian.

6 Maász, A., Kisfali, P., Horvatovich, K., Mohás, M., Markó, L., Csöngei, V., **Faragó, B**., Járomi, L., Magyari, L., Sáfrány, E., Sipeky, C., Wittmann, I. és Melegh, B.: Apolipoprotein A5 T-1131C Variant Confers Risk for Metabolic Syndrome. Pathol Oncol Res. 2007;13(3):243-7. IF=1.272

7 Komlósi, K., Talián, G.C., **Faragó, B**., Magyari, L., Cserép, V., Kovács, B., Bene, J., Havasi, V., Kiss, C.G., Czirják, L. és Melegh, B.: No influence of SLC22A4 C6607T and RUNX1 G24658C genotypic variants on the circulating carnitine ester profile in patients with rheumatoid arthritis. Clin Exp Rheumatol. 2008 Jan-Feb;26(1):61-6. IF=2.189

8 Illes Z., Safrany, E., Peterfalvi, A., Magyari, L., **Farago, B**., Pozsonyi, E., Rozsa, C., Komoly, S. és Melegh, B.: 3'UTR C2370A allele of the IL-23 receptor gene is associated with relapsing-remitting multiple sclerosis. Neurosci Lett. 2008;431:36. IF=2.200

9 **Faragó, B**. és Melegh, B.: Gilbert syndrome. Orv Hetil. 2008 Jul 6;149(27):1277-82. Review. Hungarian.

10 Szolnoki, Z., Maasz, A., Magyari, L., Horvatovich, K., **Farago, B**., Kondacs, A., Bodor, A., Hadarits, F., Orosz, P., Illes, A. és Melegh B.: Galectin-2 3279TT variant protects against the lymphotoxin-alpha 252GG genotype associated ischaemic stroke. Clin Neurol Neurosurg. 2009 Apr;111(3):227-30. IF=1.303

11 Járomi, L., Csöngei, V., Polgár, N., Szolnoki, Z., Maász, A., Horvatovich, K., **Faragó, B**., Sipeky, C., Sáfrány, E., Magyari, L., Kisfali, P., Mohás, M., Janicsek, I., Lakner, L., Melegh, B.:Functional variants of glucokinase regulatory protein and apolipoprotein A5 genes in ischemic stroke. J Mol Neurosci. 2010 May;41(1):121-8. IF=2.720

12 Csöngei, V., Járomi, L., Sáfrány, E., Sipeky, C., Magyari, L., **Faragó, B.**, Bene, J., Polgár, N., Lakner, L., Sarlós, P., Varga, M., Melegh, B.: Interaction of the major inflammatory bowel disease susceptibility alleles in Crohn's disease patients.World J Gastroenterol. 2010 Jan 14;16(2). IF=2.092

7.3. Congress abstracts and posters

1 Horvatovich, K., Magyari, L., Maasz, A., **Farago, B**., Laczy, B., Marko, L., Wittmann, I. és Melegh, B.: Association between ApoA5 T–1131C mutation and triglyceride level in Hungarian patients with metabolic syndrome and diabetes mellitus. Eur J Hum Genet 2006;14(S 1):236.

2 **Farago, B**., Talian, G.C., Maasz, A., Magyari, L., Horvatovich, K., Kovacs, B., Cserep, V., Kisfali, P., Kiss, C. és Melegh, B.: Padi4_89*G/A, padi4_90*T/C and padi4_92*G/C SNPs in the gene of the peptidylarginine deiminase citrullinating enzyme type 4 (PADI4) are not associated with rheumatoid arthritis in Hungarian patients. Eur J Hum Genet 2006;14(S 1):326.

3 Horvatovich, K., Magyari, L., Maász, A., **Faragó, B**., Laczy, B., Markó, L., Wittmann, I. és Melegh, B.: ApoA5 T-1131C mutáció és trigliceridszint közötti összefüggés vizsgálata metabolikus szindrómában és II. típusú diabetes mellitusban szenvedő betegcsoportban. Klin Kísérl Lab Med 2006;(S-32):69.

4 **Faragó, B**., Talián, C. G., Maász, A., Magyari, L., Horvatovich, K., Kovács, B., Cserép, V., Kisfali, P., Kiss, C., Czirják, L. és Melegh, B.: A peptidilarginin deiminász enzimet kódoló gén funkcionális haplotípusainak gyakorisága rheumatoid arthritis betegekben. Klin Kísérl Lab Med 2006;(S-32):101.

5 **Farago, B**., Magyari, L., Csongei, V., Jaromi, L., Safrany, E., Horvatovich, K., Sipeky, C., Maasz, A., Radics, J., Czirjak, L. és Melegh, B.: Interleukin-23 receptor 3'-UTR C2370A SNP confers risk for rheumatoid arthritis. Eur J Hum Genet 2007;15(S 1):256-257.

6 Kisfali, P., Horvatovich, K., Mohas, M., Maasz, A., Marko, L., Csongei, V., **Farago, B**., Jaromi, L., Magyari, L., Safrany, E., Sipeky, C., Wittmann, I. és Melegh, B.: Common allelic variants of APOA5 gene in the metabolic syndrome. Eur J Hum Genet 2007;15(S 1):210.

7 Csongei, V., Jaromi, L., Safrany, E., Sipeky, C., Maasz, A., Magyari, L., Horvatovich, K., **Farago, B**., Takacs, I. és Melegh, B.: Polimorphisms of the MDR1 gene in a Hungarian Roma population sample. Eur J Hum Genet 2007;15(S 1):260.

8 Safrany, E., **Farago, B**., Csongei, V., Magyari, L., Maasz, A., Sipeky, C., Jaromi, L., Horvatovich, K., Radics, J., Czirjak, L. és Melegh, B.: Interleukin-23 receptor (IL23R) gene C2370A polymorphism in scleroderma patients. Eur J Hum Genet 2007;15(S 1):256.

9 Horvatovich, K., Magyari, L., Maasz, A., Kisfali, P., Bokor, S., **Farago, B**., Csongei, V., Jaromi, L., Safrany, E., Sipeky, C., Molnar, D. és Melegh, B.: Apolipoprotein A5 T-1131C alleles in pediatric patients with obesity and metabolic syndrome. Eur J Hum Genet 2007;15(S 1):178.

10 Maasz, A., Horvatovich, K., Kisfali, P., Mohas, M., Marko, L., Csongei, V., **Farago, B**., Jaromi, L., Magyari, L., Safrany, E., Sipeky, C., Wittmann, I. és Melegh, B.: Apolipoprotein A5 T-1131C variant confers risk for metabolic syndrome. Eur J Hum Genet 2007;15(S 1):178.

11 Magyari, L., **Farago, B**., Safrany, E., Csongei, V., Horvatovich, K., Jaromi, L., Sipeky, C. és Melegh, B.: IL-23 receptor 3'UTR C2370A variant in inflammatory bowel disease: differential profile in Crohn's disease and ulcerative colitis. Eur J Hum Genet 2007;15(S 1):255.

12 Sipeky, C., Csongei, V., **Farago, B**., Horvatovich, K., Jaromi, L., Magyari, L., Safrany, E., Takacs, I. és Melegh, B.: Polymorphisms of CYP2C9 and VKORC1 genes associated with the warfarin metabolism in Hungarian Roma population. Eur J Hum Genet 2007;15(S 1):282.

13 Sipeky, C., Csöngei, V., **Faragó, B**., Horvatovich, K., Járomi, L., Maász, A., Magyari, L., Kisfali, P., Sáfrány, E., Takács I. és Melegh, B.: Vitamin K epoxide reductase haplotypes in Roma and average Hungarian samples. Clin Chem Lab Med 2008;46(8):A139.

Cumulative impact factor (not including citated abstracts): 29.130