# Analysis of interleukin-23 receptor gen polymorphisms and haplotypes in autoimmune diseases

## Doctoral (Ph.D.) theses

**Melinda Szabó, M.D.**

**Department of Medical Genetics**
**Medical School, University of Pécs**

**Project leader: Prof. Béla Melegh, M.D., Ph.D., D.Sc.**

**Pécs, 2014**

# 1. Introduction

## *1.1 The interleukin-23/interleukin-17 axis*

Cytokines are primarily intracellular connection mediator glycoproteins with low molecular weight, which play important role during immune response in the transmission and regulation of the information. These molecules are called cytokines -with other name interleukin (IL)-, which refers to the establishment of interactions between leukocytes.

The interleukin-23 (IL-23) was identified in 2000; it has a heterodimer structure, where its subunits are linked trough disulfide connections. IL-23 is composed of the p19 domain, which is typical only for IL-23; and of p40 domain, which is also a part of interleukin-12 (IL-12). In chronic inflammation, antigen-stimulated dendritic cells and macrophages produce the IL-23, which affects naive $CD4^+$ T cells and promotes the development of Th17/$Th_{IL\text{-}17}$ cells. These cells produce IL-17, which enhances T cell priming and triggers potent inflammatory responses by inducing the production of a variety of inflammatory mediators. IL-23 also acts on dendritic cells and macrophages in an autocrine manner to stimulate the production of proinflammatory cytokines, such as IL-1, IL-6, and tumor necrosis factor α (TNF-α).

Cells expressing the receptor complex built from IL-23R and IL-12Rβ1 are able to respond to IL-23. IL-23R is a trans-membrane protein built from 629 amino acids. The extracellular domain has a signal sequence, N-terminal immunoglobulin-like and two cytokine receptor domains. The human IL23R gene locates on the short arm of chromosome 1 (1p31.3). The basic form of IL-23R is coded by 11 exons; however through alternate splicing at least six different isoforms can be generated (IL-23R1-6). The most common deletions are deletion of exon 7 and/or 10. These variations result early termination of the protein leading to different forms of the extracellular domain, or to frameshift of the open reading frame, resulting different length of intracellular domains.

In 2006, genome-wide association study of Duerr et al found strong association between Crohn's disease and IL-23R polymorphisms, furthermore between IL-23R gene and polymorphism of the neighboring intergenic region of the IL-12Rß1 gene. The authors published ten different single nucleotide polymorphisms (SNP), which represented strong significant association with this type of inflammatory bowel disease. Five polymorphisms, the rs10889677 in the 3'-untranslated region (3'-UTR), the intronic rs1004819 and rs2201841, furthermore the intergenic rs11209032 and rs1495965 were found as susceptibility factors in the development of the disease. However some others were found to be protective against CD: Arg381Gln (rs11209026) in the cytoplasmic domain, the intronic rs7517847, rs10489629, rs11465804 and rs1343151 mutations.

Not much later, relationship between one of the IL23 variant (rs11805303) and CD was discovered by the Wellcome Trust Case Control Consortium analyzing the genetic background of seven autoimmune diseases. Theoretical considerations supposed the relationship of these mutations with other autoimmune diseases.

The IL-23/IL-17 axis can be influenced by many other genes. When IL-23 binds to its receptor, Janus kinases (JAK) will be activated, which phosphorilates IL-23, providing binding sites to different STAT proteins. After phosphorylation by JAK proteins, they will dimerize and translocate to the nucleus, where they stimulate the transcription of many pro-inflammatory genes (e.g. IL-17).

The JAK2 gene locates on the short arm of the chromosome 9 (9p24), and the STAT3 gene on the long arm of the chromosome 17 (17q21.31). They contain such polymorphisms (rs10758669, rs744166), which can be associated with autoimmune diseases.

## *1.2 Analyzed diseases*

### 1.2.1 Crohn's disease

Chronic inflammatory bowel diseases (IBD) are classified as Crohn's disease (CD) and ulcerative colitis (UC). Both forms may be associated with skin, eye, joint and liver alterations.

CD was originally described by Crohn, Ginzberg és Oppenheimer in 1932, as a disease localized to the ileum (ileitis terminalis, enteritis regionalis). Today, it is fact that it is a transmural and segmental process affecting the digestive tract from the mouth (buccal mucosa) until the anus. It is also characterized by recurring inflammation and unpredicted course.

In Hungary, 20-25,000 people are affected. Symptoms occur usually between 20 and 40 years affecting both sexes, but it is more common in women. Both internal and external factors play role in the development of the disease. Most important external agents are the different toxic substances, drugs, smoking, alcohol, infections and intestinal bacteria. Individual genetic characteristics, dysfunction of the immune system and various mental processes can be mentioned as internal triggers. The background of the disease is still unclear, but great importance is attached to genetic components, since some genetic predisposition may decrease the resistance of the intestinal mucosa and the dysfunction of the immune system may occur. All of these coupled with environmental factors initiate inflammatory processes.

CD is characterized by transmural and granulomatosus inflammation. It spreads segmentally where the healthy and affected bowel segments may vary. The process happens with thickening of the intestinal, stenosis and obstruction. Ulcers may penetrate; fistulas and fissures may be also formed.

Patients with CD are getting to medical care often with 4-5 years of history of abdominal pain, tactile resistance, intestinal perforation and appendicitis-like symptoms. Damage of the small intestine may lead to malabsorption. Seventy percent of the patients will need a surgery later due to different kind of complications.

Symptoms include abdominal pain, distension, diarrhea, bloody stools, formation of fistulas, fever, loss of appetite, vomiting, weight loss, growth retardation, malaise and delayed puberty. The inflammation may penetrate into the deeper layer of the bowel, leading to cohesive intestinal or colon segments. This may lead to the formation of intestinal fistula between the different abdominal organs (bladder, uterus, and vagina).

The intestinal content containing bacteria may penetrate through the fistula in to other organs, or into abdominal cavity resulting in abscesses. Formation of external fistula around the the anus is typical. The inflamed bowel is not able to absorb nutrition properly, causing often undernourishment and loss of appetite.

### 1.2.2 Ulcerative colitis

Ulcerative colitis (UC) is a chronic, inflammatory disease of the colon with unknown etiology. UC is localized exclusively for the area of the colon. The inflammation affects only the mucous membrane of the colon. The disease can be limited only for the rectum, or for the entire colon. Simultaneous inflammation of the sigmoid and the descending colon, or the sigmoid colon and the rectum are characteristics of UC.

Symptoms include abdominal pain, cramps, bloody diarrhea, constipation, fever, water and salt loss, anemia, protein and iron deficiency. Comparing to CD, only the colon is involved in UC, thus if removing of the colon is necessary, complete healing of UC is possible. (In contrast, in CD the entire digestive tract is affected from the mouth until the rectum, thus removing of the inflamed section does not cure the disease, it can return anytime.) Another important difference is based on the involvement of the intestinal segment: in CD the inflammation is limited only for the

intestinal mucosa and submucosa, but the segmental inflammation is not associated. Since the whole lumen of the bowel is affected, fistulas are not characteristics of the disorder. UC affects only the colon, so weight loss or malabsorbtion is not typical.

UC can be coupled with primer scoloritizing collangitis (cholangiocellular carcinoma as a complication!), on the other hand with duration of colitis increases the risk of colon carcinoma.

Unlike toCD, 30% of UC cases need colostomy, which can lead to full recovery.

### 1.2.3 Rheumatoid arthritis

It is a disease with unknown etiology, affecting the joins with chronic inflammatory, where genetic factors may contribute to the development of the disease. In systemic, autoimmune diseases, the body treats theirs own connective tissue elements as foreign, defending against them with inflammatory reaction. During this process, the intima of the joint thickens due to the inflammation and synovial fluid builds up. All these prevent the normal function, preventing the movement of the joint. Muscle spasms occur and at a later stage the angled form of the joint will be stabilized. During inflammation different lesions occur: degradation of the joints and surrounding areas, tendonitis is very common, rheumatoid nodules may appear in the skin, and different bone deformations may also develop. After disease manifestation, painful joint distorsion of the connective tissue or bony joint stiffness develops in the later stage.

However, in more severe form extraarticular manifestation may occur on both sides of the same affected region. Rheumatoid nodules in the skin, presence of „rheumatoid factor" (RF) in the blood as a proof, and these together with changes in the x-ray clearly demonstrate the presence of rheumatoid arthritis (RA). In 75% of patients with RA the RF is detectable, whereas in 25% of the cases not, despite in the presence of other RA specific symptoms. Some studies have shown that consumption of coffee in large amount or smoking may cause the increase of RA as the accumulation of the risk factor.

Based on epidemiological data, RA is a very common disease. In Hungary approx. 100,000 people are affected. Over 16 years of age, it can develop at any age, but most often in 40 to 50 years of age. Practically, it occurs everywhere on the world, but the incidence, the course and the severity of the disease may be different among ethnic groups, so Northern Europeans have more serious disease.

### 1.2.4. Psoriasis

Psoriasis is a genetically determined, polygenic inheritance skin disease, associated with increased proliferation of keratinocytes, paraceratosis, dermal inflammation, vascular and immunological abnormalities. It is one of the oldest known specific skin rashes, which was described already by Greek and Roman physicians as well.

The onset of the disease is influenced by climate and dietary changes, and psychological factors. As a disease, it was first described in 1808. It is a chronic, long-tern, immune-mediated disease, which tends to recurrence. In contrast to autoimmune diseases, the target antigen in not their own, instead the loss of associated flora is assumed. The estimated prevalence of psoriasis is 2% in Hungary, equally common both in women and men. It may occur at any age, but the first appearance is in the teenage with symptoms on the scalp. Ten percentages of the patients haveRA, which may lead to reduced mobility.

In the beginning, just at certain predilection areas show wax-white, stout and flaky deposits covered papules, this can remind us to dried candle wax. In other cases, especially in pediatric cases, millet size, in the later phase scaly, small and bright red papules occur on the whole body (psoriasis eruptiva, psoriasis guttata). Predilection sites of the symptoms are the surfaces areas

exposed to irritation: elbows, knees, scalp, areas under the breasts, genitofemoral bend, perianal area and the external genitalias.

### 1.2.5 Ankylosins spondylitis

The ankylosins spondilitis (Bechtrew-disease, SPA) is a chronic, inflammatory and progressive disease of the spinal and sacroiliac joints, especially characterized with calcification of the joints and ligaments. As a consequence of these results a rigid spine. The peripheral joins are affected in the one third of the cases.

The prevalence of the disease is 0.1-1.4%. Both sexes are afftected; however it is more common in man, than women. In 80% of the cases, the symptoms manifest before 30 years of age, and develop only in 5% of the cases over 45 years of age. Initial symptoms are usually low back or thoracic pain, which occurs with variable strength mostly at night, and worsens gradually over the time.

As the joint surfaces die due to the inflammation, the joins calcificates, the vertebrae begin to grow together, and beaklike lumps are formed on their edges. SPA may affect the entire body. Approximately 25% of the patients show some kind of non-musculoskeletal complications, such as iritis, aortic valve insufficiency, circulatory failure or fibrosis of the upper lung lobes.

# 2. Aims

The aim of our experiments was to analyze the IL-23R gene polymorphisms and their haplotypes in certain autoimmune diseases in the Hungarian population:

To study the genetic variants of IL-23R gene rs1004819, rs7517847, rs7530511, rs2201841, rs1343151 and rs10889677 in rheumatoid arthritis, Crohn’s disease and in ankylosis spondylitis.

We also wanted to examine the role of the following variants: rs1884444, rs11805303, rs7517847, rs2201841, rs10889677 and rs11209032 in the two aspects of inflammatory bowel disease and psoriasis.

Our goal was to analyze the role of these variants and their common haplotypes in the development of the above mentioned diseases.

Another goal was to analyze two known polymorphisms of JAK2 and STAT3 (rs10758669 and rs744166) which influences the IL-23 activity, in Hungarian patients with CD and UC.

# 3. Patients and methods

## *3.1 Study populations*

For the haplotype analysis of IL-23R gene we used 190 (mean age: 39.0 ± 14.0 years) and 199 (mean age: 38.8 ± 14.1 year) patients with CD. All samples originated from the Biobank governed by the University of Pecs, as a part of the National Biobank.

DNA samples of the total of 282 patients with UC (mean age: 46.7 ± 16.1 years) originated from the Biobank governed by he University of Pecs, as well as the total of 396 DNA samples of patients with RA (mean age:54.9 ± 14.7 years).

The SPA group contained a total of 206 patients (mean age: 40.7 ± 15.4 years). These samples originated from the National Institute of Rheumatology and Physiotherapy. By comparison, DNA samples of 255 healthy controls (mean age: 45.0 ± 10.9 years) were used.

For the analysis of the IL-23R gene polymorphisms, a total of 263 DNA samples from patients with psoriasis were used (mean age: 47.5 ± 12.3 years). The collection of the blood samples was done at the University Clinic of Dermatology and Allergology, Szeged. As a control, DNA samples from 189 clinically healthy patients were used (mean age: 44.6 ± 12.0 years).

For the analysis of JAK2 and STAT3 gene polymorphisms DNA samples of a total of 309 patients with CD (38.7 ± 0.80 years) and 307 with UC (45.2 ± 0.90 years) were used, and as a control DNS samples of 496 healthy adults were used (46.5 ± 0.80 years).

DNA samples of the control persons originated in all cases from the Biobank of the University of Pecs, Department of Medical Genetics. Participants of these trials gave their written consent for genetic testing in advance. Each study was conducted in accordance with the permission of the Ethics Committee (ETT TUKEB), and the 1964 Declaration of Helsinki principles.

## *3.2 Molecular biological methods*

The genomic DNA was isolated from peripheral blood leukocytes with routine salting out method. For genotyping of the variants PCR-RFLP method was applied. Restriction fragments were separated by gel electrophoresis containing ethidium bromide and visualized by UV transillumination.

Specific primers and restriction endonucleases used to the amplification, and the enzyme cleavage patterns are shown in **Table 1**. In the case of rs2201841, mismatch base was applied in primer design to create an artificial cleavage site (underlined in the sequence).

## *3.3 Statistical analysis*

To the evaluation of the relationship between the diseases, the genetic variants and the haplotypes, $\chi^2$-test and regression analysis was used. Analyses were carried out SPSS 11.5 package for Windows. Haploview 4.1 was used to test linkage disequilibrium. Haplotype frequencies were estimated using PHASE version 2.1.

***Table 1***. Primer sequences, length of PCR products, restriction endonucleases and enzyme specific cleavage patterns used in the experiments

| | Forward primer | Reverse primer | Length of the PCR product (bp) | Restriction endonuclease | Enzyme cleavage pattern of the homozygote common allele (bp) | Enzyme cleavage pattern of the heterozygote genotype (bp) | Enzyme cleavage pattern of the homozygote rare allele (bp) |
|---|---|---|---|---|---|---|---|
| **IL23R rs1884444** | CAGTCTTTTCCTGCTTCCAGACAT | AATAAAATCATACTCTTGCCAATGGCCC | 509 | PscI | 191+318 | 28+191+290+318 | 28+191+290 |
| **IL23R rs1004819** | GCATTCTAGGACCGTTTTGG | ATCTGGTGGAAATATGTGAAACCTA | 270 | TaaI | 13+71+185 | 13+71+185+257 | 13+257 |
| **IL23R rs11805303** | TCTTCCCAGTCTCCAGTGTG | CCGAACAATTTTTGTTTCCC | 373 | MnlI | 39+136+198 | 39+136+198+237 | 136+237 |
| **IL23R rs7517847** | AAACATTGACATTCCCTTCATAC | GAAATGAGTCACCAATAATCCAC | 530 | BseMII | 29+91+410 | 29+91+410+501 | 29+501 |
| **IL23R rs7530511** | TACCCATCCATTTTAGGTTAAAGAA | GTCTTGAAGTCCTGACCTAAGGTAATC | 614 | HphI | 51+134+429 | 51+134+185+429 | 185+429 |
| **IL23R rs2201841** | GGCAAAAGGGAATTGAGAGG | GGCCTATGATTATGCTTTTTCCTG | 420 | HpyF3I | 163+257 | 25+163+232+257 | 25+163+232 |
| **IL23R rs1343151** | ACAAATTTTGACTTGAATGTTCTTTTCC | AAATGAGCAAAGAATTGCCCT | 479 | Hpy188I | 27+154+298 | 27+154+181+298 | 181+298 |
| **IL23R rs10889677** | ATCGTGAATGAGGAGTTGCC | TGTGCCTGTATGTGTGACCA | 470 | MnlI | 61+185+224 | 61+185+224+285 | 185+285 |
| **IL23R rs11209032** | TTGTTACTGGAGTTAAACCTCTTGC | AGGAATAATTGCTGAGATGCAATG | 265 | BseMI | 24+67+174 | 24+67+174+242 | 24+242 |
| **JAK2 rs10758669** | GATCTGTTCACTGGCAATATCTTTT | TCTTAAGGAGGGACATAAATGTGAG | 445 | BseDI | 53+392 | 53+100+292+392 | 53+100+292 |
| **STAT3 rs744166** | GCTGTAATGTCTTGAGGGAATCAA | GTAGAGACTGTACAAGGAGACCACG | 497 | HindIII | 23+109+365 | 23+109+365+388 | 109+388 |

# 4. Results

## *4.1 Analysis of rheumatoid arthritis, Crohn's disease and ankylosis spondilitis*

Six polymorphisms of the IL-23R gene were analyzed in patients with RA (n=396), CD (n=190 and SPA (n=206) and compared with healthy controls (n=182). All genotype and allele frequencies were in Hardy-Weinberg equilibrium in all four groups. Genotype and allele frequencies are shown in **Table 2.**

***Table 2.*** Case-control genotypes and minor allele frequencies (MAF) in RA, CD, SPA and control groups

| | | RA (n=396) | CD (n=190) | SPA (n=206) | Controls (n=182) |
|---|---|---|---|---|---|
| **rs1004819** | **GG** | 169 (42.7) | 66 (34.7) | 74 (35.9) | 91 (50.0) |
| | **GA** | 180 (45.5) | 102 (53.7) | 110 (53.4) | 73 (40.1) |
| | **AA** | 47 (11.9) | 22 (11.6) | 22 (10.7) | 18 (9.89) |
| | **MAF** | 0.35 | 0.38* | 0.37* | 0.30 |
| **rs7517847** | **TT** | 150 (37.9) | 71 (37.4) | 67 (32.5) | 57 (31.3) |
| | **TG** | 182 (46.0) | 102 (53.7) | 115 (55.8) | 99 (54.4) |
| | **GG** | 64 (16.2) | 17 (8.95) | 24 (11.7) | 26 (14.3) |
| | **MAF** | 0.39 | 0.36 | 0.40 | 0.41 |
| **rs7530511** | **CC** | 295 (74.5) | 145 (76.3) | 159 (77.2) | 132 (72.5) |
| | **CT** | 90 (22.7) | 39 (20.5) | 43 (20.9) | 46 (25.5) |
| | **TT** | 11 (2.77) | 6 (3.16) | 4 (1.94) | 4 (2.20) |
| | **MAF** | 0.14 | 0.13 | 0.12 | 0.15 |
| **rs2201841** | **TT** | 167 (42.2) | 75 (39.5) | 91 (44.2) | 82 (45.1) |
| | **TC** | 175 (44.2) | 90 (47.4) | 89 (43.2) | 89 (48.9) |
| | **CC** | 54 (13.6)* | 25 (13.2)* | 26 (12.6)* | 11 (6.04) |
| | **MAF** | 0.36 | 0.37 | 0.34 | 0.31 |
| **rs1343151** | **GG** | 174 (43.9) | 97 (51.1) | 102 (49.5) | 70 (38.5) |
| | **GA** | 190 (48.0) | 82 (43.2) | 96 (46.6) | 97 (53.3) |
| | **AA** | 32 (8.08) | 11 (5.79) | 3 (3.88) | 15 (8.24) |
| | **MAF** | 0.32 | 0.27* | 0.27* | 0.35 |
| **rs10889677** | **CC** | 173 (43.7) | 75 (39.5) | 88 (42.7) | 77 (42.3) |
| | **CA** | 174 (43.9) | 92 (48.4) | 94 (45.6) | 94 (51.6) |
| | **AA** | 49 (12.4)* | 23 (12.1)* | 24 (11.7)* | 11 (6.04) |
| | **MAF** | 0.34 | 0.36 | 0.35 | 0.32 |

** $p<0.05$ vs. controls*

Significant different was observed in the allele frequencies of two SNPs in CD and SPA groups compared to control subjects. The most significant different was found in the CD group in the case of the variant rs1004819 (p=0.015; OR=1.46; 95% CI: 1.08-1.98), similarly to this the polymorphism rs1343151 was significant too (p=0.027, OR=0.70, 95% CI: 0.52-0.97). Minor allele frequency of rs1004819 showed significant different in SPA patients (p=0.029; OR=1.40; 95% CI: 1.03-1.89), as well as in the case of rs1343151 (p=0.020; OR=0.70; 95% CI: 0.51-0.94), versus to controls. The homozygote form of the variant rs2201841 was found as a susceptibility factor in all three analyzed diseases. In the case of CD patients: p=0.013; OR=2.35; 95% CI: 1.20-4.60; at SPA patients the level of significance was a bit lower (p=0.031; OR=2.24; 95% CI: 1.08-4.68), whereas in RA patients it was stronger (p=0.009; OR=2.45; 95% CI: 1.25-4.81). Minor allele carriers of the variant rs10889677 in the 3' UTR region increased the odds ratio (>2x) in all analyzed disease groups. In RA: p=0.023; OR=2.20; 95% CI: 1.11-4.33; in SPA p=0.047; OR=2.10; 95% CI: 1.01-4.53; in CD p=0.032; OR=2.10; 95% CI: 1.06-4.14.

The major haplotypes (ht) are shown in **Table 3.** A total of 4 haplotypes showed association with one of the analyzed diseases. Haplotype frequencies are shown in **Table 4.** Two haplotypes were found as susceptibility and two protective at one the analyzed disease. The ht1 haplotype decreased the odds ratio RA (p=0.029; OR=0.66; 95% CI: 0.46-0.96). The ht2 occurs significantly more times in patients with RA or SPA, compared to healthy control subjects.

***Table 3.*** Major haplotypes (ht) created by the examined IL-23R variants

| **Haplotípus** | **rs1004819** | **rs7517847** | **rs7530511** | **rs2201841** | **rs1343151** | **rs10889677** |
|---|---|---|---|---|---|---|
| **ht1** | G | G | C | T | G | C |
| **ht2** | A | T | C | C | G | A |
| **ht3** | G | T | C | T | G | C |
| **ht4** | G | G | C | T | A | C |
| **ht5** | A | T | C | C | A | A |
| **ht6** | G | T | T | T | G | C |
| **ht7** | G | T | T | T | A | C |

In the case of RA it caused 1.44 higher odds ratio (p=0.019; OR=1.44; 95% CI: 1.06-1.96), whereas in SPA the odds ratio was much higher (p=0.006; OR=1.78; 95% CI: 1.18-2.68). The haplotype ht4 showed to be as a susceptibility factor in RA (p=0.037; OR=1.38; 95% CI: 1.10-1.96), however at the other two diseases it was protective, but the results were statistically not significant. The haplotype ht7 showed strong protective effect in all three diseases, reducing their odds ration to 1/3. In patients with RA p=0,005; OR=0.26; 95% CI: 0.10-0.67; in CD p=0.009; OR=0.37; 95% CI: 0.17-0.78; in SPA p=0.004; OR=0.38; 95% CI: 0.19-0.74.

Several strong significant differences were found between RA and the other two diseases in the distribution of the major haplotypes (**Table 4**). The haplotype ht1 occurs more often in RA than in SPA (p<0.001), and also compared to CD (p<0.001). The haplotype ht2 occurred less times in CD than in RA (p<0.001), whereas ht3 showed increased level compared to RA (p=0.002). The haplotypes ht4 and ht6 were significantly more frequent in RA, than in CD (p<0.001 respectively p=0.027) and in SPA (p<.001in both two haplotypes). The haplotype ht5 occurred rarely in patients with RA, compared to patients with CD or SPA (p<0.001).

***Table 4.*** Haplotype frequencies of the examined IL-23R variants

| | **RA (%)** | **CD (%)** | **SPA (%)** | **Kontroll (%)** |
|---|---|---|---|---|
| **ht1** | 12.2* | 18.8§ | 23.1§ | 20.6 |
| **ht2** | 23.5* | 15.1§ | 19.7* | 13.3 |
| **ht3** | 9.17 | 15.3§ | 13.5 | 12.1 |
| **ht4** | 16.1* | 4.97§ | 6.62§ | 9.22 |
| **ht5** | 3.95 | 9.98§ | 7.68§ | 7.12 |
| **ht6** | 9.17 | 5.99§ | 4.66§ | 5.50 |
| **ht7** | 1.25* | 2.46* | 2.21* | 4.72 |

**p<0.05 vs. controls; §p<0.05 vs. RA patients*

## *4.2 Analysis of ulcerative colitis, Crohn's disease and psoriasis*

Six polymorphisms of the IL-23R gene was analyzed in patient with UC (n=282) CD (n=199) and psoriasis (n=263, and compared to healthy control subjects (n=253). All genotype and allele frequencies were in Hardy-Weinberg equilibrium in all four groups. Genotypes and allele frequencies are shown in **Table 5.** The TT genotype of the variant rs1884444 showed strong susceptibility in patients with UC (p=0.001; OR=3.13; 95% CI: 1.60-6.13), or psoriasis (p=0.005; OR=2.68; 95% CI: 1.35-5.35). The GG genotype of rs7517547 is protective against the development of CD (p=0.017; OR=0.48; 95% CI: 0.27-0.88), whereas the homozygote variants of rs2201841 and rs10889677 were found as a susceptibility factors (p=0.007; OR=2.43; 95% CI: 1.27-4.62 respectively p=0.016; OR=2.28; 95% CI: 1.17-4.45). The last mentioned SNP showed association with psoriasis (p=0.041; OR=1.97; 95% CI: 1.03-3.76).

A total of four haplotypes showed association with one of the analyzed diseases. The major haplotype frequencies are shown in **Table 6.** Three haplotypes conferred as a susceptibility factor in the development in one the analyzed diseases. The haplotype ht5 occurred significantly more often in patients with UC, compared to healthy subjects. (p=0.003; OR=2.50; 95% CI: 1.36-4.58). Two haplotypes, the ht6 and ht8 showed strong predisposing affect for the development of CD; the haplotype ht6 showed the strongest significance level compared to all diseases (ht6 p<0.001; OR=3.04; 95% CI: 1.69-5.46 and ht8 p=0.020; OR=2.62; 95% CI: 1.16-5.89).

Analysis of the major haplotypes resulted several strong associations between UC and the other two diseases (**Table 7.**). The haplotype ht1 was more common in UC, than in CD (p<0.001) and psoriasis (p=0.045). The ht6 haplotype occurred with less frequency in UC, than in CD (p=0.002), whereas the haplotype ht4 occurred more frequent in UC, than in CD (p=0.021). Haplotype frequency of ht6 showed significant different between patients with psoriasis and CD (p=0.006).

***Table 5.*** Case-control genotypes and minor allele frequencies (MAF) in CD, UC, psoriasis and in control groups

| | | CD (n=199) | UC (n=282) | Psoriasis (n=263) | Control (n=253) |
|---|---|---|---|---|---|
| **rs1884444** | **GG** | 60 (30.2) | 68 (24.1) | 75 (28.5) | 61 (24.1) |
| | **GT** | 133 (66.8) | 176 (62.4) | 157 (59.7) | 180 (71.1) |
| | **TT** | 6 (3.00) | 38 (13.5)* | 31 (11.8)* | 12 (4.74) |
| | **MAF** | 0.36 | 0.45 | 0.42 | 0.40 |
| **rs11805303** | **CC** | 80 (40.2) | 135 (47.9) | 111 (42.2) | 129 (51.0) |
| | **CT** | 99 (49.7) | 124 (44.0) | 126 (47.9) | 102 (40.3) |
| | **TT** | 20 (10.1) | 23 (8.16) | 26 (9.98) | 22 (8.70) |
| | **MAF** | 0.35 | 0.30 | 0.34 | 0.29 |
| **rs7517547** | **TT** | 72 (36.2) | 115 (40.8) | 89 (33.8) | 74 (29.2) |
| | **TG** | 110 (55.3) | 122 (43.3) | 128 (48.7) | 138 (54.5) |
| | **GG** | 17 (8.54)* | 45 (16.0) | 46 (17.5) | 41 (16.2) |
| | **MAF** | 0.36 | 0.38 | 0.42 | 0.43 |
| **rs2201841** | **TT** | 77 (38.7) | 132 (46.8) | 89 (33.8) | 123 (48.6) |
| | **TC** | 94 (47.2) | 129 (45.7) | 128 (48.7) | 114 (45.1) |
| | **CC** | 28 (14.1)* | 21 (7.45) | 46 (17.5) | 16 (6.32) |
| | **MAF** | 0.38 | 0.30 | 0.42 | 0.29 |
| **rs10889677** | **CC** | 77 (38.7) | 128 (45.4) | 128 (48.5) | 117 (46.2) |
| | **CA** | 97 (48.7) | 133 (47.2) | 106 (40.3) | 121 (47.8) |
| | **AA** | 25 (12.6)* | 21 (7.45) | 29 (11.0)* | 15 (5.93) |
| | **MAF** | 0.37 | 0.24 | 0.31 | 0.30 |
| **rs11209032** | **GG** | 69 (34.7) | 134 (47.5) | 127 (48.3) | 109 (43.1) |
| | **GA** | 107 (53.8) | 122 (43.3) | 107 (40.7) | 120 (47.4) |
| | **AA** | 23 (11.6) | 26 (9.22) | 29 (11.0) | 24 (9.49) |
| | **MAF** | 0.38 | 0.31 | 0.31 | 0.33 |

** $p<0.05$ vs. controls*

***Table 6.*** Major haplotypes (ht) created by the examined IL-23R variants

| | **rs1884444** | **rs11805303** | **rs7517847** | **rs2201841** | **rs10889677** | **rs11209032** |
|---|---|---|---|---|---|---|
| **ht1** | G | C | G | T | C | G |
| **ht2** | G | C | T | T | C | G |
| **ht3** | T | T | T | C | A | A |
| **ht4** | T | C | T | T | C | G |
| **ht5** | T | C | G | T | C | G |
| **ht6** | G | T | T | C | A | A |
| **ht7** | G | C | G | T | C | A |
| **ht8** | T | T | T | C | A | G |

***Table 7***. Haplotype frequencies of the examined IL-23R variants

| | **CD (%)** | **UC (%)** | **Psoriasis (%)** | **Control (%)** |
|---|---|---|---|---|
| **ht1** | 21.1 | 21.0 | 24.7 | 24.7 |
| **ht2** | 19.9 | 20.1 | 17.0 | 17.4 |
| **ht3** | 12.8 | 15.1 | 14.3 | 12.9 |
| **ht4** | 5.50 | 9.47$^{\S}$ | 9.33 | 8.42 |
| **ht5** | 4.83 | 9.72*$^{\S\#}$ | 5.91 | 6.55 |
| **ht6** | 10.4* | 6.14$^{\S}$ | 6.71$^{\S}$ | 4.99 |
| **ht7** | 2.32 | 1.27 | 1.68 | 2.44 |
| **ht8** | 3.62* | 3.41 | 2.07 | 1.67 |

**p<0.05 vs. controls; $^{\S}$p<0.05 vs. CD; $^{\#}$p<0.05vs psoriasis*

## *4.3 Polymorphisms of JAK2 and STAT3, as a possible escalations*

One to one polymorphisms of the genes JAK2 and STAT3 were tested in patients with CD (n=308) and UC (n=307). The results of these were compared to 496 healthy control subjects. All genotype and allele frequencies were in Hardy-Weinberg equilibrium in all three groups. Genotypes and allele frequencies are shown in **Table 8.**

The TT homozygote genotype showed strong susceptibility in patients with UC (p=0.009; OR=1.48; 95% CI: 1.10-1.99) compared to control subjects. Frequency of the T minor allele showed significant different compared to controls. Thus, we can conclude that this polymorphism of STAT3 is a risk factor for ulcerative colitis. Regarding to patients with CD this association was not detectable.

In the case of the variant rs10758669 (JAK2) we could not detect any significant different compared to patients with CD and UC, when they were compared to controls (neither in the case of homozygotes, nor in the allele frequencies).

***Table 8.*** Case-control genotypes and minor allele frequencies (MAF) in CD, UC and in control groups

| | | **CD (n=309)** | **UC (n=307)** | **Controls (n=496)** |
|---|---|---|---|---|
| **STAT3 rs744166** | **CC** | 33 (10.7%) | 32 (10.4%) | 72 (14.5%) |
| | **CT** | 163 (52.7%) | 133 (43.3%) | 243 (49.1%) |
| | **TT** | 113 (36.6%) | 142 (46.3%) | 181 (36.4%) |
| | **CT+TT** | 276 (89.3%) | 275 (89.6%) | 424 (85.3%) |
| | **MAF** | 62.9% | 67.9% | 60.98% |
| **JAK2 rs10758669** | **AA** | 112 (36.2%) | 127 (41.4%) | 203 (40.8%) |
| | **AC** | 159 (51.5%) | 142 (46.2%) | 243 (49.1%) |
| | **CC** | 38 (12.3%) | 38 (12.4%) | 50 (10.1%) |
| | **AC+CC** | 197 (63.8%) | 180 (58.6%) | 288 (57.9%) |
| | **MAF** | 38.0% | 35.50% | 34.41% |

** p<0.05 vs. controls*

# 5. Discussion of the results and conclusions

The IL-23 -as a member of the IL-12 cytokine family- plays important role in various inflammatory processes, therefore modifications in this gene and the consequent changes in the structure, lifetime and function of the protein may affect certain autoimmune functions and the development or severity of these diseases.

Our goal was to analyze in Hungarian patients and healthy subjects the different gene polymorphisms and haplotypes of IL-23 in autoimmune diseases and their possible effect on the development of the disease.

Association of SPA and the IL-23R gene was shown at first time by the genome-wide linkage analysis of the Wellcome Trust Case Control Consortium. A total of eight SNPS of the IL-23R gene were analyzed (rs11209026, rs1004819, rs10489629, rs11465804, rs1343151, rs10889677, rs11209032, rs1495965), which showed association with the disease. The strongest association was found with the variant rs11209032. Our results were supported by a study on Spanish population, where the rs11209026 (Arg381Gln) and rs1343151 was found to be protective against the development of the disease. Similar results were found in British population, where similarly to the above mentioned study population the variant rs11209032 showed the strongest association with SPA, whereas in Portuguese patients the SNP rs1004819 was proved as a highest risk factor.

In the international literature the association of rheumatoid arthritis and IL-23R gene is somewhat contradictory. Our research group found association between two variants (rs10889677 and rs2201841) and RA in Hungarian population, however other studies not. Recently, in a large population study an association was found between the SNP rs1343151 SNP and the disease. These observations encouraged us to extend our analyses for these polymorphisms, and to carry out haplotype analysis. In the haplotype analysis we included those SNPs which were found to be significantly associated with RA and three other variants which showed association with other autoimmune diseases.

An association was found between the polymorphism rs1343151 and CD and compared to SPA too; however that kind of association was not observable in patients with RA. Similarly to this, any association was found with the other three SNPs (rs1004819, rs7517847, rs7530511). The four major IL-23R haplotypes showed association with one of the tested diseases. From this haplotypes two (ht1 and ht4) showed association only with RA. This fact appears to be substantial and draws attention to how important is not just to analyze the individual polymorphisms, but their combined haplotypes may help for better understanding of the diseases. These were supported by the fact, while the seven major haplotypes showed any significant difference between CD and SPA, six showed significant difference between RA and/or CD and SPA. Surprisingly this difference was more pronounced than in healthy controls and in analyzed diseases. In addition the frequency of ht4 was increased in RA compared to controls, whereas in the other two diseases reduced, concluding the direction of the association is the opposite in RA compared to CD and SPA. These results are accordance with the observations of Holliss-Moffatt, who found the same in the rare A allele of the polymorphism rs1343151. This allele was also included in ht4.

The IL-23R gene and psoriasis was associated at first time by Cargill and coworkers. Based on these results, the presence of most common haplotype defined by the C allele of rs7530511 and the G allele of rs11209026 G is significantly higher in patients with psoriasis, compared to healthy control subjects. Association between the latter SNP, respectively the haplotype combined from the two mutations and psoriasis was confirmed by numerous studies. Association was found between rs2201841 and psoriasis too.

Results testing the association of UC and IL-23R gene are somewhat contradictory. Duerr and coworkers showed at the first time between the disease and this gene. Shortly after, a study in Chilean patients did not find any association, while in Spanish patients only a trend was observable,

which was statistically not significant. Another study on European inflammatory bowel disease patients found association between the polymorphism Arg381Gln and CD, but also in UC.

Overall, these observations led us to examine the relationship of IL-23 polymorphisms and haplotypes in Hungarian UC patients. Genotyping of the same polymorphisms was performed in CD, which is the other type of IBD, in which was shown several times the important role of IL-23R. In our study we involved psoriasis, which is symptomatic completely different.

Association was found between the variant rs1884444, psoriasis and UC, but not with CD. The three other polymorphims (rs7517847, rs2201841, rs100889677) showed association with CD and psoriasis, but not with UC. From the analyzed eight major haplotypes in eight was found association with one of the tested diseases, compared to healthy controls. One of the haplotypes, the ht5 showed association only with UC, indicating how important is not to analyze the individual polymorphisms alone; their combined haplotypes needs to be analyzed too. The distribution of the analyzed haplotypes showed large differences in all three diseases. In psoriasis in comparison to CD, statistically significant difference was found in one haplotype (ht6), but interestingly between the two forms of IBD the three haplotypes differed. The haplotype ht5 occurred more frequent in UC, compared to the other diseases. The increases frequency of ht6 seems to be as n isolation marker in CD. All these suggest, not only the individual polymorphisms and their haplotypes may differ, but apparently similar diseases may differ in terms of their genetic background.

It is clear; the haplogroup analysis shows further progress in understanding the role of IL-23 in autoimmune diseases. However, it is clear hat many other factors can affect the effect of IL-23, not just its polymorphisms. As a taken out example, DNA samples of CD and UC patients were analyzed for JAK2 rs10758669 and STAT3 rs744166 polymorphims. In this regard, we found the STAT3 rs744166 is a susceptibility factor in Hungarian UC patients, compared to control subjects both in T allele frequency and the level of TT homozygotes.

Our results contribute to accurate knowledge and understanding of the background of some autoimmune diseases, which is essential for an early detection of the disease. All these contribute to the identification of ne therapeutic targets. Of course, there is a need to identify additional genetic variants and careful genotype-phenotype analyses, because this is the way to the effective prevention in the increasingly expanding knowledge, and in better management of autoimmune diseases.

# 6. Summary of the results

1. Studying of the IL-23R gene variants we found, that the presence of the variants rs2201841 and rs10889677 in homozygote forms increase the risk of RA, CD and SPA. Homozygote form of the variant rs1004819 increases the risk of CD and SPA.

2. The GGCTGC, ATCCGA, GGCTAC and GTTTAC haplotypes of the variants rs1004819, rs7517847, rs7530511, rs2201841, rs1343151 and rs10889677 are associated with RA, and the GTTTAC haplotype with CD and SPA.

3. The GGCTAC haplotype frequency of the variants rs1004819, rs7517847, rs7530511, rs2201841, rs1343151 and rs10889677 significantly distinct in the RA group compared to healthy controls, patients with CD or SPA.
4. The homozygous form of the polymorphism rs1884444 was significantly increased in UC and psoriasis, compared to controls.

5. The GTTCAA and TTTCAG haplotypes of the variants rs1884444, rs11805303, rs7517847, rs2201841, rs10889677 and rs11209032 are more common in CD, and the TCGTCG haplotype in UC.

6. The TCGTCG haplotype frequency of the variants rs1884444, rs11805303, rs7517847, rs2201841, rs10889677 and rs11209032 significantly distinct in patients with UC compared to healthy controls, patents with CD disease or psoriasis.

7. Analysis of the variants JAK2 rs10758669 and STAT3 rs744166 showed that homozygote TT form of STAT3 rs744166 and the T allele frequency significantly increased in patients with UC compared to controls.

# 7. Publications

## 7.1. Publications supporting the dissertation

1. **Szabo M**, Safrany E, Pazar B, Melegh BI, Kisfali P, Poor G, Figler M, Szekanecz Z, Czirjak L, Melegh B. Marked diversity of IL23R gene haplotype variants in rheumatoid arthritis comparing with Crohn's disease and ankylosing spondylitis. Mol Biol Rep. 2013 Jan;40(1):359-63.
**IF: 2.506**

2. Safrany E, **Szabo M**, Szell M, Kemeny L, Sumegi K, Melegh BI, Magyari L, Matyas P, Figler M, Weber A, Tulassay Z, Melegh B. Difference of interleukin-23 receptor gene haplotype variants in ulcerative colitis compared to Crohn's disease and psoriasis. Inflamm Res. 2013 Feb;62(2):195-200. **IF: 1.964**

3. Polgar N, Csongei V, **Szabo M**, Zambo V, Melegh BI, Sumegi K, Nagy G, Tulassay Z, Melegh B. Investigation of JAK2, STAT3 and CCR6 polymorphisms and their gene-gene interactions in inflammatory bowel disease. Int J Immunogenet. 2012 Jun;39(3):247-52. **IF: 1.355**

## 7.2. Other publications

1. Sipeky C, Csongei V, Jaromi L, Safrany E, Maasz A, Takacs I, Beres J, Fodor L, **Szabo M**, Melegh B. Genetic variability and haplotype profile of MDR1 (ABCB1) in Roma and Hungarian population samples with a review of the literature. Drug Metab Pharmacokinet. 2011;26(2):206-15. **IF: 2.321**

2. Polgár N, Járomi L, Csöngei V, Maász A, Sipeky C, Sáfrány E, **Szabó M**, Melegh B. Triglyceride level modifying functional variants of GALTN2 and MLXIPL in patients with ischaemic stroke. Eur J Neurol. 2010 Aug;17(8):1033-9. **IF: 3.765**

3. Sipeky C, Lakner L, **Szabo M**, Takacs I, Tamasi V, Polgar N, Falus A, Melegh B. Interethnic differences of CYP2C9 alleles in healthy Hungarian and Roma population samples: relationship to worldwide allelic frequencies. Blood Cells Mol Dis. 2009 Nov-Dec;43(3):239-42. **IF: 2.901**

4. Sipeky C, Csongei V, Jaromi L, Safrany E, Polgar N, Lakner L, **Szabo M**, Takacs I, Melegh B. Vitamin K epoxide reductase complex 1 (VKORC1) haplotypes in healthy Hungarian and Roma population samples. Pharmacogenomics. 2009 Jun;10(6):1025-32. **IF: 3.893**

## 7.3. Citable abstracts

1. Janicsek I, Polgar N, Csongei V, **Szabo M**, Zambo V, Melegh B: Associations of STAT3, JAK2 and CCR6 polymorphisms with ulcerative colitis and Crohn's diasease. Eur J Hum Genet. 2011; 19, Suppl 2: 262.

2. Sipeky C, Maasz A, Saghy E, Tarlos G, **Szabo M**, Takacs I, Melegh B: Polymorphisms of thiopurine S-methyltransferase (TPMT) gene in the average Roma and Hungarian population samples. Eur J Hum Genet. 2011; 19, Suppl 2: 344.

3. Safrany E, **Szabo M**, Melegh B: Interleukin-23 receptor gene haplotypes in diseases known to associate with individual interleukin-23 receptor gene mutations. Eur J Hum Genet. 2010; 18, Suppl 1: 257.

4. Sipeky C, Safrany E, Csongei V, Jaromi L, **Szabo M**, Kisfali P, Maasz A, Polgar N, Bene J, Takacs I, Melegh B: Haplotype profile of multidrug resistance 1 (MDR1/ABCB1) gene in the average Hungarian and Roma population samples. Eur J Hum Genet. 2010; 18, Suppl 1: 259.

5. Sipeky C, Safrany E, Csongei V, Jaromi L, Kisfali P, Maasz A, Polgar N, Bene J, Takacs I, **Szabo M**, Melegh B: Comparison of VKORC1 haplotype profile and CYP2C9 polymorphisms as determinants of coumarin dose in Hungarian and Roma population samples. Eur J Hum Genet. 2009; 17, Suppl 1: 280.

**Cumulative impact factors (without citable abstracts): 18.705**

# 8. Acknowledgements

The research, which serves as a basis of the doctoral dissertation, was carried out at the Department of Medical Genetics.

First of all, I would like to express my greatest thanks to my supervisor, Professor Béla Melegh for rendering it possible for me to join the Ph.D. program; for his confidence and his supervision of my work, for the leadership and useful advices.

I would like to thank for each colleague in the laboratory especially for Dr. Judit Bene Berenténé, Dr. Lili Magyari, Dr. Enikő Sáfrány and Dr. Luca Járomi who greatly contributed to the successful work.

Finally, I express my warmest thanks to my family for their love and encouraging support.