NKX2-3 and IRGM variants are associated with disease susceptibility to IBD in Eastern European patients

Nora Meggyesi, Lajos S Kiss, Magdalena Koszarska, Martin Bortlik, Dana Duricova, Laszlo Lakatos, Tamas Molnar, Martin Leniček, Libor Vítek, Istvan Altorjay, Maria Papp, Zsolt Tulassay, Pal Miheller, Janos Papp, Attila Tordai, Hajnalka Andrikovics, Milan Lukas, Peter Laszlo Lakatos

Abstract

AIM: To investigate variants of immunity-related GTPase family M (IRGM) and NKX2-3 genes and genotype-phenotype in Eastern European patients with inflammatory bowel disease (IBD).

METHODS: We analyzed 1707 Hungarian and Czech subjects with Crohn’s disease (CD) (n = 810, age: 37.1 ± 12.6 years, duration: 10.7 ± 8.4 years) and ulcerative colitis (UC) (n = 428, age: 43.7 ± 15.0 years, duration: 12.6 ± 9.9 years), as well as 469 healthy controls. IRGM rs13361189, NKX2-3 rs10883365 and ECM1 rs13294 polymorphisms were tested by LightCycler allele discrimination. Detailed clinical phenotypes were determined by reviewing the medical charts.

RESULTS: NKX2-3 rs10883365 variant allele was associated with increased risk for CD (P = 0.009, OR = 1.24, 95% CI = 1.06-1.48) and UC (P = 0.001, OR = 1.36, 95% CI = 1.13-1.63), whereas variant IRGM allele increased risk for CD (P = 0.029, OR = 1.36, 95% CI = 1.03-1.79). In contrast, ECM1 rs13294 was not associated with CD or UC. In CD, the variant IRGM allele was associated with a colon-only location (P = 0.009, OR = 1.62, 95% CI = 1.07-2.44), whereas in UC, the ECM1 variant was associated with cutaneous manifestations (P = 0.002, OR = 3.36, 95% CI = 1.48-7.63). Variant alleles did not predict resistance to steroids or azathioprine, efficacy of infliximab, or need for surgery.

CONCLUSION: NKX2-3 and IRGM are susceptibility loci...
INTRODUCTION

Inflammatory bowel diseases (IBDs) are multifactorial with both environmental and genetic components; the latter displaying heterogeneity in terms of disease presentation as well as response to treatment\(^a\). Crohn’s disease (CD) has a strong genetic component, and at least nine susceptibility loci have been identified\(^b\). The first, and most consistently replicated critical mutations have been found in the NOD2/CARD15 gene on chromosome 16 (IBD1). Other loci encode genes that are involved in a number of homeostatic mechanisms: innate pattern recognition receptors (including NOD2/CARD15, TLR4 and CARD9), differentiation of Th17-lymphocytes (IL-23R, JAK2, STAT3, CCR6 and ICOSLG), autophagy [ATG16L1, immunity-related GTPase family M (IRGM) and LRRK2], maintenance of epithelial barrier integrity (IBD5, DLG5, PTGER4, ITLN1, DMBT1 and XBPI), and orchestration of the secondary immune response (HLA-region, TNFSF15/TL1A, IRE6, PTPN2, PTPN22, NFX2-3, IL-12B, IL-18RAP and MST1)\(^c\). As the incidence of IBD is rapidly increasing in some parts of Eastern Europe\(^b\), it is of great importance to study social and environmental, as well as host-genetic factors that might underlie this trend.

The IRGM gene is located on chromosome 5q33.1, and encodes an autophagy-inducing protein and belongs to the immunity-related guanosine triphosphatases (IRGs), also known as p47 guanosine triphosphatases. IRGs play an important role in host defense against intracellular pathogens\(^d\). Genetic association of the autophagy genes, autophagy-related 16-like 1 (ATG16L1) and IRGM has been suggested in adult-onset CD by genome-wide association scan studies (GWASs), but not in UC\(^e-f,9\). Although a meta-analysis of the three index GWASs has implicated a single nucleotide polymorphism (SNP; rs11747270) based on imputed data with an OR of 1.33\(^i\), other studies including a new meta-analysis\(^j\) have confirmed the IRGM signal on rs13361189 and rs4958847 immediately flanking IRGM.

Recently, the rs13361189 variant, upstream of IRGM, has been shown to be in perfect linkage disequilibrium with a 20-kb deletion polymorphism that affects the expression of IRGM (and cellular autophagy) in a tissue-specific manner\(^k\). These results suggest that the CD association at IRGM arises from an alteration in IRGM regulation that affects the efficacy of autophagy, and identifies rs13361189 and its strongly correlated neighbors at the 5’ end of IRGM, including the 20-kb deletion polymorphism as a likely causal variant.

Association between the rs10883365 variant of NK2 transcription factor related, locus 3 (NFX2-3) gene on chromosome 10q24.2 and susceptibility of CD has been reported in Western Europe, including The Wellcome Trust Case Control Consortium (WTCCC) GWA and a replication study from The Netherlands, and in Japan, with an OR of 1.2-1.6\(^l-m\). In addition, a modest association (\(P = 3.3 \times 10^{-4}\) in the ulcerative colitis panel and \(P = 2.4 \times 10^{-6}\) using the expanded WTCCC control panel) has also been reported between rs10883365 and UC in a non-synonymous SNP scan by Fisher et al\(^n\). NFX2-3-deficient mice show severe defects in gut development; primarily in the epithelium of the small intestine\(^o\). In addition, the lymphoid organs of these mice, including the spleen and Peyer’s patches, have abnormal tissue architecture and abnormalities in the migration and segregation of B and T cells\(^p\), however, the exact mechanism is unknown.

ECM1, on chromosome 1q21.2, is also a plausible candidate gene for UC; it encodes extracellular matrix protein 1, a glycoprotein expressed in small and large intestine, and it interacts with the basement membrane and inhibits matrix metalloproteinase\(^q\), Notably, ECM1 strongly activates nuclear factor-kB signaling, a key immune regulator. Expression is upregulated in colorectal cancer and metastases, which implicates ECM1 in epithelial-stromal interaction\(^r\). Of note, the WTCCC observed modest association between these ECM1 SNPs and ankylosing spondylitis, a related inflammatory disorder (\(P = 0.0041 \) and 0.0044), respectively\(^s\). More recently, an association between the rs3737240 and rs13294 variants of ECM1 and UC susceptibility was reported in a GWA by Fisher et al\(^t\), with an OR of 1.3-1.4 for the homozygous carriage of the variant allele. rs3737240 and rs13294 encode substitutions T130M and G2908: Thr130, residing within a collagen IV binding domain, is conserved in primates, whereas Gly290 is not.

Finally, data about the pharmacogenetics of IBD are still limited. Resistance to steroids is associated with high expression of β-glucocorticoid receptors (hGRβ) and over-expression of MDR1 has been found in patients who fail steroid therapy and require surgery\(^u\). In contrast, according to our group’s earlier results, the presence of the DLG5 Arg30Gln variant allele, but not variants in ABCG2 or MDR1 genes, predicted resistance to steroids\(^v\). Vermeire et al\(^w\) have reported a lack of association between the
presence of NOD2 mutations and response to infliximab therapy, and later, the same group reported an association between the presence of Fas ligand -843C/T variant and response to infliximab\(^{23}\).

In light of the lack of data in Eastern European countries, our aim was to investigate the prevalence of IRGM rs13361189, NKK2-3 rs10883365 and ECM1 rs13294 variants in large independent cohorts of Czech and Hungarian IBD patients. We also aimed to investigate a possible association between genotype and clinical phenotype, need for surgery, and response to medical therapy.

**MATERIALS AND METHODS**

**Study population**

One thousand seven hundred and seven unrelated IBD patients (CD: \(n = 810\), age: 37.1 ± 12.6 years, duration: 10.7 ± 8.4 years and UC: \(n = 428\), age: 43.7 ± 15.0 years, duration: 12.6 ± 9.9 years) and 469 healthy subjects (blood donors) from Hungary and the Czech Republic were investigated. The clinical data of the CD and UC patients are presented in Table 1. A detailed clinical phenotype was available in 789 CD and 422 UC patients.

The diagnosis was based on the Lennard-Jones criteria\(^{24}\): age; age at onset; presence of extraintestinal manifestations [arthritis: peripheral and axial; ocular manifestations: conjunctivitis, uveitis and iridocyclitis; skin lesions: erythema nodosum and pyoderma gangrenosum; and hepatic manifestations: primary sclerosing cholangitis (PSC)]; frequency of flare-ups (frequent relapses: > 1 clinical relapse/year); therapeutic effectiveness (e.g. steroid and/or immunosuppressive use, steroid resistance); need for surgery (resections); the presence of familial IBD; smoking habits; and in CD, perianal involvement, were investigated by reviewing the medical charts by the physician and completing a questionnaire. The disease phenotype (age at onset, duration, location and behavior) was determined according to the Montreal classification\(^{25}\).

The control group for mutation analysis consisted of 469 age- and sex-matched healthy blood donors (male/female: 251/218, age: 40.5 ± 11.5 years old). Control subjects did not have any gastrointestinal and/or liver diseases and were selected from consecutive blood donors in Budapest, Veszprem, Debrecen and Prague. The study protocol was approved by the Ethical and Science Committee of the Ministry of Health. Each patient was informed of the nature of the study and gave signed informed consent.

**Genotyping and DNA isolation**

Genomic DNA was isolated from whole blood according to the manufacturers’ description with High Pure PCR Template preparation Kit (Roche, Budaors, Hungary). Amplification primers and hybridization probes were designed by the LightCycler Probe Design software (Roche Diagnostics). All oligonucleotides were synthesized by VBC Biotech (Vienna, Austria). The following amplification primers and hybridization probes were used for genotyping IRGM: IRGM-LCF: 5'-ATGGACAGTG-CAGTACCCGTGCAC-3', IRGM-LCR: 5'-CTCTTTTAC-CTTGTACCTTGC-3', IRGM-SENS: 5'-GAAAATCGGATG-640-TGCTCAGCGGGTACAGTTTAGAAAGGGAA-Phosphate-3', IRGM-ANC: 5'-LC Red 640-TGCTCAGCGGGTACAGTTTAGAAAGGGAA-Phosphate-3', IRGM-SENS: 5'-GAAATCGGATG-TATATATTAGTAGACCC-Fluorescein-3'.

The amplification primers and hybridization probes used for genotyping ECM1 were: ECM1-LCF: 5'-ACCCACCACTTTGTT-3', ECM1-LCR: 5'-TGCTTG-GTGAAGAACACTCTTGGT-3', ECM1-ANC: 5'-LC Red 640-TGCTCAGCGGGTACAGTTTAGAAAGGGAA-Phosphate-3', ECM1-SENS: 5'-TGACTGACCGATGTCAAT-Fluorescein-3'.

The amplification primers and hybridization probes used for genotyping of NKK2-3 were: NKK2-2-LCF:
Association between NKX2-3, IRGM and ECM1 and disease susceptibility

All investigated polymorphisms were in Hardy-Weinberg equilibrium \((P = 0.38-0.95)\). The success rate of the genotyping assays was 98%-99%. The genotype and allele frequencies are presented in Table 2. NKX2-3 rs10883365 variant allele was associated with increased risk for CD \((P = 0.018 \text{ after Bonferroni correction, } OR = 1.24, 95\% \text{ CI } = 1.06-1.48)\) and UC \((P = 0.003 \text{ after Bonferroni correction, } OR = 1.36, 95\% \text{ CI } = 1.13-1.63)\), whereas the variant IRGM allele increased the risk of CD \((P = 0.04 \text{ after Bonferroni correction, } OR = 1.42, 95\% \text{ CI } = 1.16-1.72)\) and UC \((P = 0.04 \text{ after Bonferroni correction, } OR = 1.36, 95\% \text{ CI } = 1.06-1.48)\) and UC \((P = 0.003 \text{ after Bonferroni correction, } OR = 1.36, 95\% \text{ CI } = 1.06-1.48)\). The association between NKX2-3 rs10883365 variant and IBD was also significant in the genotypic and dominant models. A similar trend was noted between IRGM and CD. In contrast, ECM1 rs13294 was not associated with either CD or UC. The combination of IRGM carrier/NKX2-3 homozygote genotype was significantly higher in CD compared to the controls \((7.2\% \text{ vs } 2.1\%, P = 0.0001 \text{ after Bonferroni correction, } OR = 3.55, 95\% \text{ CI } = 1.80-7.01)\).

**Role of NKX2-3, IRGM and ECM1 variants in predicting sex, familial disease, age at onset, disease location, phenotype or extraintestinal manifestations**

In CD, presence of the variant IRGM allele was associated with colon-only location \((25\% \text{ vs } 5\%, P = 0.04)\). No other significant associations were found in either CD or UC patients (data not shown).

**Association between NKX2-3, IRGM and ECM1 and response to medical therapy or need for surgery**

We also investigated the association between the NKX2-3, IRGM and ECM1 variants and the response to steroids,
infliximab or azathioprine or the need for surgery in patients with CD. Two hundred and sixty-three unrelated CD patients were treated with anti-tumor necrosis factor (TNF): 259 received infliximab and four, adalimumab, two of whom were treated after secondary loss of response to infliximab therapy [5 mg/kg at weeks 0, 2 and 6; partial response (assessed at week 12) to infliximab induction therapy] (data not shown). Similarly, no association was found between the presence of the above variants and short-term infliximab therapy. Overall, there was no association between either of the variants and steroid use/resistance, azathioprine use or need for surgery in patients with UC (data not shown).

DISCUSSION

This is the first report on the prevalence of the IRGM rs13361189, NXK2-3 rs10883365 and ECM1 rs13294 variants in large, independent IBD cohorts from Eastern Europe. The variant NXK2-3 allele conferred a risk in both UC and CD, whereas the IRGM rs13361189 variant allele was associated with increased risk for CD. In addition, the IRGM variant was associated with disease location in CD.

The genotype and allele frequencies for IRGM and NXK2-3 variants in CD, UC and controls reported in the present study were in line with most previous reports in Caucasian populations, whereas the frequency of the variant ECM1 rs13294 allele was approximately 10% higher in both IBD and controls, compared to that reported in previous studies [1,2,3,13,15,26].

We confirmed that the IRGM rs13361189 variant was associated with disease susceptibility in CD in Eastern European populations. The allele OR of 1.36 (95% CI = 1.03-1.79) was in the range reported in Caucasian populations (OR = 1.38-1.56) [7,13,14] and in a recent meta-analysis (OR = 1.34) [15]. Similar genotype distributions were observed in both the Czech and Hungarian cohort, however, due to the low variant allele frequency, the difference became significant only in the combined analysis. A similar tendency (P = 0.06) was observed in the genotypeic and dominant model. In contrast, in the present study we failed to replicate the weak association between IRGM and UC (rs13361189 P = 0.0069, pooled OR = 1.16; rs4958847 P = 0.014, pooled OR = 1.13) that was reported recently in a Spanish meta-analysis [19]. Of note, the present study was underpowered to detect such small
differences, however, there was even no trend for a difference between UC and controls.

In addition, in an Italian study\cite{27}, the rs4958847 polymorphism was associated with fistulizing behavior ($P = 0.037$, OR = 1.54, CI = 1.02-2.31) and perianal fistulas ($P = 0.045$, OR = 1.55, CI = 1.01-2.38) in a logistic regression analysis. This was later partially confirmed by the Leuven group. Henckaerts et al.\cite{28} have reported in a very elegant study a significant association between the IRGM rs4958847 variant, U7 gene desert rs12704036 T-allele, NOD2/CARD15 mutation, ileal involvement at diagnosis, male sex, and time to development of non-perianal fistulas in a Cox regression analysis. In a further study from New Zealand\cite{29}, rs13361189 variant increased the risk for ileal CD in 507 CD patients and 576 controls. The OR in ileal CD was 1.92 (95% CI = 1.27-2.96). Moreover, Peterson et al.\cite{30} have suggested an association between IRGM and pediatric disease onset (<17 years) of CD in a North American cohort. However, only one of the two IRGM variants studied (rs13361189) was weakly associated with CD in their study (uncorrected $P = 0.03$). In the present study, we could not confirm the association between age at onset and the presence of the IRGM variant, in accordance with Canadian, Italian and Scottish results\cite{31,32}. In contrast, we found an association between IRGM carriage and disease location in CD; colonic location was significantly more common in carriers of the variant allele (OR = 1.62, 95% CI = 1.07-2.44). Of note, however, the rs4958847 variant was not investigated in the present study.

The association between the rs10883365 variant of NKKX2-3 gene and susceptibility of IBD was first reported in CD in Caucasian patients and in a Japanese study with an OR of 1.2-1.6\cite{33,34}. In addition, using the UC panel and the expanded WTCCC control panel, Fisher et al.\cite{35} have reported a modest association ($P = 3.3 \times 10^{-4}$ and $P = 2.4 \times 10^{-7}$) between rs10883365 and UC. In the present study, we confirmed these findings; the rs10883365 variant was associated with UC and CD susceptibility in the allelic and genotypic models with an OR of 1.53 and 1.24, respectively. Based on our data, the association between NKKX2-3 and IBD is stronger in UC compared to CD, at least in patients from Eastern Europe. In a recent Dutch study, the association between the rs10883365 variant of NKKX2-3 and CD was linked to smoking status, with the risk being more pronounced in active and passive smokers\cite{36}. In the present study, however, we did not find an association between NKKX2-3 and smoking status, and similarly, no genotype-phenotype associations were found.

In addition, Weersma et al.\cite{37} have created genetic risk profiles by combining the presence of variant alleles in IL23R, ATG16L1, IRGM, NKKX2-3, 1q24, 5p13, HERC2, CCNY, 1q21 and NOD2/CARD15, and the number of risk alleles was associated with gradually increased risk for CD. Similarly, in the present study, the combination of IRGM carrier/NKKX2-3 homozygote genotype was associated with an increased risk for CD, with a much higher OR compared to either of the variant alleles alone. Recently, an association between ECM1 rs3737240 and rs13294 variants and UC has been reported in a GWAS by Fisher et al.\cite{38}, with an OR of 1.3-1.4 for the homozygous carriage of the variant allele. The association for the rs13294 variant has recently been confirmed in a Dutch study, with an OR of 1.24 in patients with UC\cite{39}. In contrast, no association was found in CD\cite{40}, even though a different SNP (rs11205387) was investigated. Although the present study was powered to confirm a difference with an OR of 1.3-1.5 with adequate statistical power, we could not confirm an association between the ECM1 rs13294 variant and either UC or CD. The accidental association between homozygous carriage of the ECM1 variant allele and cutaneous manifestations ($P = 0.002$, OR = 3.36, 95% CI = 1.48-7.63, Figure 1B) requires further confirmation.

Theoretically, through influencing inflammatory responses, autophagy and epithelial-stromal interactions, polymorphisms in the above genes might be of potential importance in altering the efficacy of anti-inflammatory therapy and thereby the need for surgery. However, in the present study, none of the variant alleles was associated with the response to either steroid or infliximab therapy, or the need for surgery (resection in CD or colectomy in UC) in IBD.

In conclusion, NKKX2-3 and IRGM are susceptibility loci for IBD in Eastern European patients. None of the variants investigated were associated with the need for surgery or efficacy of medical therapy. Further studies are needed to confirm the reported phenotype-genotype associations found in this study.

**ACKNOWLEDGMENTS**

Further Hungarian IBD Study Group members are: Simon Fischer, 1st Department of Medicine, Semmelweis University, Budapest; Zsuzsanna Erdelyi, Gabor Mester, Tunde Pandur, 1st Department of Medicine, Csonlony F County Hospital, Veszprem; Agota Kovacs, Laszlo Bene, 1st Department of Medicine, Erzsebet Hospital, Budapest; Ferenc Nagy, Klaudia Farkas, 1st Department of Medicine, University of Szeged, Szeged; Karoly Palatka, Zsuzsanna Vitalis, 2nd Department of Medicine, University of Debrecen, Debrecen; Laszlo Herszenyi, 2nd Department of Medicine, Semmelweis University, Budapest.

**COMMENTS**

**Background**

Sequence variants in the autophagy gene immunity-related GTPase family M (IRGM) and NKKX2-3 have been reported to contribute to Crohn’s disease (CD) susceptibility, whereas ECM1 contributes to ulcerative colitis (UC) in genome-wide association scans in North America and Western Europe.

**Research frontiers**

There is a lack of data in Eastern European countries, therefore, our aim was to investigate the prevalence of IRGM rs13361189, NKKX2-3 rs10883365 and ECM1 rs13294 variants. In addition, the possible association between genotype and clinical phenotype and the need for surgery is conflictive, and the association between the above genetic variants and response to medical therapy was not investigated.

**Innovations and breakthroughs**

In the present study, the authors showed in two well-characterized, independent
CD cohorts with strict clinical follow-up, that NKX2-3 rs18833365 variant allele was associated with increased susceptibility to CD and UC, whereas variant IRGM allele increased the risk for CD only. Our data suggest that the variant IRGM allele is associated with colon-only location, whereas in UC, the ECM1 variant was associated with cutaneous manifestations. None of the variants predicted resistance to steroids and azathioprine, efficacy of infliximab, or need for surgery.

**Application**

The new data presented from Eastern Europe might contribute to better understanding of genetic or environmental differences between populations and the association of those differences with disease susceptibility and phenotype.

**Terminology**

Vienna-Montreal classification: classification systems of disease phenotypes in CD. The classification assesses the age at presentation, disease location and disease behavior. The IRGM gene encodes an autophagy-inducing protein and plays an important role in host defense against intracellular pathogens. NKX-3 is a member of the NKX family of homeodomain-containing transcription factors, which are implicated in many aspects of cell type specification and maintenance of differentiated tissue functions. ECM1 encodes extracellular matrix protein 1, a glycoprotein expressed in small and large intestine. Notably, ECM1 strongly activates nuclear-S signaling, a key immune regulator.

**Peer review**

The need to extend studies performed in the other parts of the world to Eastern Europe is valid, and provides a point for understanding subtle genetic or environmental differences between populations and the association of those differences with disease.

**REFERENCES**


12. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007; 447: 661-678


20. Fischer S, Lakatos PL, Lakatos L, Kovacs A, Molnár T, Alterjagy I, Papp M, Szilvási A, Tulasay Z, Ozsvótit J, Papp J, De-menter CE, Abbondandi A, Andersson H. ATP-binding cassette transporter ABCG2 (BCRP) and ABCB1 (MDR1) variants are not associated with disease susceptibility, disease phenotype response to medical therapy or need for surgery in
Megyesi N et al. NKX-IRGM in inflammatory bowel disease

Hungarian patients with inflammatory bowel diseases. *Scand J Gastroenterol* 2007; **42**: 726-733


S-Editor Cheng JX L-Editor Kerr C E-Editor Ma WH