Biochimica et Biophysica Acta 1820 (2012) 1512–1518

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagen

Non-synonymous single nucleotide polymorphisms in genes for immunoregulatory galectins: Association of galectin-8 (F19Y) occurrence with autoimmune diseases in a Caucasian population

Zsuzsanna Pál [a,b,1], Péter Antal [c,1], Sanjeev Kumar Srivastava [b], Gábor Hullám [c], Ágnes F. Semsei [b], János Gál [d], Mihály Svébis [d], Györgyi Soós [d], Csaba Szalai [b], Sabine André [e], Elena Gordeeva [f], György Nagy [b,g], Herbert Kaltner [e], Nicolai V. Bovin [f], Mária Judit Molnár [a], András Falus [b,h], Hans-Joachim Gabius [e,1], Edit Irén Buzás [b,*,1]

[a] Center for Molecular Neurology, Department of Neurology, Semmelweis University, Budapest, Hungary
[b] Department of Genetics, Cell- and Immunobiology, Semmelweis University, Budapest, Hungary
[c] Department of Measurement and Information Systems, Budapest University of Technology and Economics, Budapest, Hungary
[d] Department of Rheumatology, County Hospital, Kecskemét, Hungary
[e] Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University Munich, Munich, Germany
[f] Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry RAS, Moscow, Russia
[g] Department of Rheumatology, Semmelweis University, Frankel Leó u. 54, Budapest, Hungary
[h] Research Group for Inflammation Biology and Immunogenomics, Hungarian Academy of Sciences, Budapest, Hungary

ARTICLE INFO

Article history:
Received 2 January 2012
Received in revised form 29 May 2012
Accepted 31 May 2012
Available online 7 June 2012

Keywords:
Allele
Autoimmunity
Lectin
Neoglycoconjugate
Sialylation

ABSTRACT

Background: Galectins are potent immune regulators, with galectin-8 acting as a pro-apoptotic effector on synovial fluid cells and thymocytes and stimulator on T-cells. To set a proof-of-principle example for risk assessment in autoimmunity, and for a mutation affecting physiological galectin sensor functions, a polymorphism in the coding region of the *galectin-8* gene (rs2737713; F19Y) was studied for its association with two autoimmune disorders, i.e. rheumatoid arthritis and myasthenia gravis.
Methods: A case–control analysis and a related quantitative trait-association study were performed to investigate the association of this polymorphism in patients (myasthenia gravis 149, rheumatoid arthritis 214 and 134 as primary and repetitive cohorts, respectively) and 365 ethnically matched (Caucasian) healthy controls. Distribution was also investigated in patients grouped according to their antibody status and age at disease onset. Comparative testing for lectin activity was carried out in ELISA/ELLA-based binding tests with both wild-type and F19Y mutant galectin-8 from peripheral blood mononuclear cell lysates of healthy individuals with different genotypes as well as with recombinant wild-type and F19Y mutant galectin-8 proteins.
Results: A strong association was found for rheumatoid arthritis, and a mild one with myasthenia gravis. Furthermore, the presence of the sequence deviation also correlated with age at disease onset in the case of rheumatoid arthritis. The F19Y substitution did not appear to affect carbohydrate binding in solid-phase assays markedly.
General significance: This is the first report of an association between a galectin-based polymorphism leading to a mutant protein and autoimmune diseases, with evidence for antagonistic pleiotropy.

1. Introduction

During the past years numerous susceptibility alleles have been described for a series of human autoimmune diseases. It has become clear that such alleles can often occur in more than one autoimmune disorder [1]. As a consequence, their identification can help to explain why certain autoimmune diseases such as myasthenia gravis (MG), a rare autoimmune disease affecting neuromuscular transmission and rheumatoid arthritis (RA), a highly disabling disease that causes destruction of the joints, are often diagnosed in the same patient [2]. Of course, the development of these diseases will be associated with

Abbreviations: AChR, acetylcholine receptor; ANOVA, analysis of variances; BMLA, Bayesian multilevel analysis; BSA, bovine serum albumin; CCP, cyclic citrullinated peptide; CD, cluster of differentiation; CI, confidence interval; ELISA, enzyme-linked immunosorbent assay; ELLA, enzyme-linked lectin assay; Exp(B), exponentiation of the B value; HLA, human leukocyte antigen; HWE, Hardy–Weinberg equilibrium; LGALS8, human galectin-8; MG, myasthenia gravis; OD, optical density; OR, odds ratio; PBMC, peripheral blood mononuclear cell; PBS, phosphate-buffered saline; RA, rheumatoid arthritis; SD, standard deviation; SE, standard error of mean; SNP, single nucleotide polymorphism

* Corresponding author at: Department of Genetics, Cell and Immunobiology, Semmelweis University, Nagyvárad tér 4, H-1089 Budapest, Hungary. Tel./fax: +36 30 7444918.
E-mail address: buzedi@dgci.sote.hu (E.I. Buzás).
[1] Z.P. and P.A. as well as H.-J.G. and E.I.B. contributed equally to this work.

0304-4165/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbagen.2012.05.015

a complex pattern of inherited traits, with an array of genes being involved in their development and progression [3,4]. Among the genetic factors, HLA-DRB1, for instance, accounts for approximately 30% of the genetic susceptibility in the case of RA [5], while a series of HLA associations were delineated among MG patients [6]. Using the sophisticated tools of molecular genetics, genome-wide association studies are unveiling "common" non-HLA autoimmunity genes, giving current research to define molecular mechanisms of disease manifestation a clear direction [7]. Obviously, it is a promising aim to identify new susceptibility genes and risk-associated single nucleotide polymorphisms (SNPs). Based on the increasing body of evidence for a role of glycan–protein (lectin) interactions in inflammatory pathways [8–10], we here test the hypothesis of a respective potential for a family of adhesion/growth-regulatory proteins, also known to be strongly immunomodulatory, i.e. the galectins [11,12].

These endogenous lectins are known to target distinct glycan and peptide motifs (signals) and hereby trigger biosignaling for growth control/mediator release or modulate adhesion/migration of cells [10–12]. What attracted our attention was that the tandem-repeat-type galectin-8 (for a review on its structure and physiological roles, please see [13]) is an inducer of apoptosis of synovial fluid cells, an activity neutralized by a distinct glycan in the rheumatoid version of CD44 (CD44vRA) [14], and that auto-antibodies against this lectin are present in the sera of patients with systemic lupus erythematosus (23%), rheumatoid arthritis (16%) and septicaemia (20%) [15]. By virtue of its remarkable affinity to α2,3-sialylated N-glycans, distinct sulfated N-acetyllactosamine forms and histo-blood group B epitope, assigned to either the N- or C-domain [16–22], this galectin could have a special function in autoimmune regulation. Such a role is intimated by its T cell-stimulatory activity, assumed to enhance otherwise borderline immune responses [23], acting in a concentration-dependent manner on proliferation and as co-stimulator of antigen-specific responses [23,24]. In order to reveal a hypothetically postulated relationship, we focused on an SNP in the coding region of the *LGALS8* gene of patients with RA and MG, in relation to healthy controls. The examined non-synonymous mutation affects amino acid 19 in exon 1, causing a substitution of phenylalanine (F) to tyrosine (Y). Since the case study of human galectin-1 mutant with serine replacing cystein at position 2 has taught the lesson that a seemingly minor change can have structurally long-range effects with an impact on entropic contributions in the thermodynamic balance sheet of ligand binding in the β-sandwich fold [25], the quest to define clinical associations, if successful, will give ensuing structural work a clear direction.

2. Patients and methods

2.1. Patients and controls

We recruited 149 patients with MG (122 women, 27 men, mean age: 50 ± 13.2 years (range:18–78 years), mean age at disease onset: 36.2 ± 13.4 years) from the NEPSYBANK (Hungarian Neurological and Psychiatric Biobank) [26]; 214 RA patients (180 women, 34 men, mean age: 56.7 ± 9.4 years (range: 26–85 years), mean age at disease onset: 53 ± 9 years) treated at the Department of Rheumatology, County Hospital, Kecskemét, Hungary, and 365 ethnically matched Caucasian controls (263 women, 102 men, mean age: 37.5 ± 13.7 years (range: 18–83 years)), without known autoimmune disease. These were blood donors or enrolled during routine check-ups. To add strength to the results, a separate, independent repetitive RA cohort was additionally analyzed in this study, which was composed of 134 patients (26 men, 108 women, mean age: 64.2 ± 11.1 years (range: 29–84)) treated at the Department of Rheumatology, Semmelweis University. Diagnosis in the case of MG was based on standard criteria including symptoms of fluctuating muscle weakness supported by an electromyographic pattern of neuromuscular transmission dysfunction by repetitive stimulation. All RA patients met the American College of Rheumatology/European League Against Rheumatism criteria for rheumatoid arthritis set in 2010 [27].

Determination of age at onset was based on the patient's recollection and/or previous clinical documentation. All patients and controls gave explicit written informed consent. This study was carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans, and it was approved by the local ethical committees.

2.2. Prediction of functional SNPs in the human galectin-8 gene

We searched for known *galectin-8* SNPs in exons using dbSNP from the www.ncbi.nlm.nih.gov and HapMap (www.Hapmap.org) websites. By now, 4 missense SNPs have been identified in the *GALS8* gene. Out of these, only rs1041935, rs1041937 and the examined rs2737713 encode changes in putatively functionally relevant sites (localized in the carbohydrate-recognition domains), while the rs2243525 will affect the linker of the long isoform which is the minor species due to prevalent representation of the short linker caused by alternative splicing removing the respective sequence stretch. The rs2737713 A > T SNP of the *LGALS8* gene was selected as it was the only sequence deviation according to the annotation of dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/), which may affect both ligand binding (via a long-range effect first defined for human galectins in the case of the C2S mutant of galectin-1 [25]) and inter-domain interactions. In the Caucasian population, "A" is the major and "T" is the minor allele. Polymorphism of the contingent T allele in the gene causes a missense mutation from F to Y at amino acid 19 (F19Y substitution) in the N-terminal domain of the galectin-8 protein. We tested whether this SNP could contribute to autoimmune disease susceptibility by investigating its distribution in Caucasian MG and RA patients.

2.3. Determination of anti-acetyl-choline receptor (anti-AChR) antibody concentration

The level of AChR-specific antibody was measured by radioimmunoprecipitation (DRG Diagnostics, Marburg, Germany) according to the manufacturer's instructions. Following incubation of the patients' sera with ^{125}I-labeled α-bungarotoxin, complexes of labeled receptor and antibody were immunoprecipitated with anti-human IgG and radioactivity quantitated in a γ-counter. The measured radioactivity was directly proportional to the concentration of anti-AChR autoantibodies of the sample.

2.4. Determination of anti-cyclic citrullinated peptide (anti-CCP) antibodies

Level of serum antibodies reactive to CCP was measured with a commercial ELISA (Eurodiagnostica AB, Malmö, Sweden) according to the manufacturer's instructions.

2.5. DNA isolation and polymorphism analysis

Genomic DNA was extracted from peripheral blood cells using the QIAamp DNA blood kit (Qiagen, Valencia, CA, USA). Genotypes were determined using pre-developed TaqMan Allelic Discrimination Assay (Assay ID: C_1653103_10, the context sequence used was (VIC/FAM) AATTTTTTTTCTTAGGTAATCCCGT[A/T]TGTTGGCACCATTCCTGATCAGCTG; Applied Biosystems, Foster City, CA, USA). Briefly, polymerase chain reactions (PCR) were carried out in a 96-well format with the GeneAmp PCR System 7000 (Applied Biosystems) with aliquots of mixtures consisting of 10 ng of genomic DNA, 5 μl of TaqMan Universal PCR Master Mix, 0.25 μl of 40× assay mix, and double-distilled water up to 10 μl final volume. Thermal cycle conditions were as follows: denaturation at 95 °C for 10 min, followed by 40 cycles of denaturation at 92 °C for 15 s, and annealing and extension at 60 °C for 1 min. After

PCR the endpoint fluorescence intensity was read in each well. The allele-specific intensity data of each plate were analyzed using the ABI PRISM 7000 software (Applied Biosystems) to automatically determine the genotype of each sample.

2.6. Isolation of peripheral blood mononuclear cells (PBMCs)

PBMCs were isolated with Histopaque 1191/1117 (Sigma-Aldrich, Munich, Germany) gradient from healthy controls previously genotyped for the *LGALS8* rs2737713 SNP. Cells were lysed in cell lysis buffer (ProteoJet Mammalian Cell lysis reagent; Fermentas, Burlington, ON, Canada) supplemented with the protease inhibitors leupeptin and aprotinin (at 1 μM), 200 μM phenylmethylsulfonyl fluoride and 100 mM sodium orthovanadate (all from Sigma-Aldrich) just before use.

2.7. Protein determination

The protein concentration in each sample was determined with a Micro BCA Protein Assay kit (Thermo Scientific, Rockford, IL, USA) using serum albumin supplied with the kit as a standard. Measurements were done according to the protocol of the manufacturer. Absorbance was measured with the BCA protein measurement protocol on a Nanodrop 1000 v.3.7 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA).

2.8. Determination of galectin-8 concentration by ELISA

ELISA plates (Nunc, Wiesbaden, Germany) were coated with 0.25 μg/well polyclonal anti-human galectin-8 antibody (R&D Systems, Minneapolis, MN, USA) diluted in bicarbonate buffer (pH 9.6). Plates were incubated for 2 h at 37 °C and then overnight at room temperature. Blocking was carried out using 1% PBS-BSA for 2 h at 37 °C. 30 μg of protein was used from the cell lysates in a total volume of 100 μl diluted in PBS and was incubated overnight at 37 °C. After washing three times with PBS containing 0.1% Tween 20 (washing buffer), biotin-labeled polyclonal goat anti-human galectin-8 antibody (R&D Systems) was added at a 1:500 dilution for 2 h at 37 °C. After washing three times with washing buffer, a solution containing streptavidin-HRP (Invitrogen Corporation, Carlsbad, CA, USA) at a 1:2500 dilution was used for 2 h at 37 °C. Thereafter, o-phenylenediamine (Sigma-Aldrich) and 0.33% hydrogen peroxide in PBS was applied, and the absorbance after signal development, was detected at 492 nm. For each plate standard curves were obtained using known amounts of human galectin-8 (standard linker length), obtained by recombinant production and purification, rigorously controlled by mass spectrometry and two-dimensioned electrophoresis for purity as well as solid-phase and cell assays for activity [15,28].

2.9. Binding assays by enzyme-linked immunosorbent assay (ELISA)/ enzyme-linked lectin assay (ELLA)

Polyacrylamide-based neoglycoconjugates (Lectinity, Moscow, Russia) were absorbed to the surface of microtiter plate wells (Nunc, Wiesbaden, Germany) using bicarbonate buffer (pH 9.6) at 0.5 μg/well for testing cell lysates and 0.25 μg/well for testing the recombinant proteins, as routinely done with glycoproteins in solid-phase lectin assays [29]. Plates were incubated for 2 h at 37 °C and then overnight at room temperature. Blocking of remaining sites for protein binding on the plastic surface was carried out using 1% PBS–BSA for 2 h at 37 °C. In testing of cell lysates (obtained from 16 AA and 9 TT genotyped healthy individuals), we applied 70 μg of protein from cell lysates in a total volume of 100 μl diluted in PBS supplemented with protease inhibitors to the wells, and incubated the plates overnight at 37 °C. For the testing of recombinant proteins, two-fold serial dilutions of known amounts of protein were incubated in plates overnight at 37 °C. After washing, biotin-conjugated polyclonal goat anti-human galectin-8 antibody (R&D Systems) was used at a 1:500 dilution for 2 h at 37 °C. Streptavidin-HRP (Invitrogen Corporation) was added at a 1:2500 dilution for 2 h at 37 °C and absorbance was measured after signal development in a spectrophotometer as described above.

Alternatively, biotinylated galectin-8 (N-terminal domain and full-length protein) was used as a sensor as described [29]. The mutant proteins were obtained by entering the T/A exchange into the galectin-8-encoding cDNA (AF342815) using a modified QuickChange™ site-directed mutagenesis procedure (Agilent Technologies Stratagene, Santa Clara, CA, USA) and ensuing recombinant production as done with wild-type proteins [15,19,28].

Data processing of ELISA assays was carried out using Graphpad Prism version 4 software (Graphpad Software Inc., San Diego, CA, USA). Standard curves were obtained using known concentrations of the standard isoform of recombinant human galectin-8 in the cell-lysate assays (total galectin-8 determination and binding assays). Deviation from normal distribution was evaluated with the Kolmogorov–Smirnoff test. Comparison between the different genotyped healthy individuals was done using a t-test or Mann–Whitney test, where appropriate, with the SPSS 17.0 software (SPSS Inc., Chicago, IL, USA).

For comparison of carbohydrate-binding capacities of the wild-type and the F19Y mutant N-terminal domains of galectin-8, ODs from known amounts of serially diluted proteins were compared in three independent experiments. Statistical analysis of data from assays with neoglycoconjugates was performed using two-way repeated measures ANOVA with a Holm–Sidak post-hoc test (SigmaStat for windows version 3.5 software, Systat Software Inc.).

2.10. Statistical analysis of clinical data

The SNP was assessed for deviation from Hardy–Weinberg equilibrium (HWE) in all examined populations (Table 1) by using an online software (http://ihg2.helmholtz-muenchen.de/cgi-bin/hw/hwa1.pl) as described [30].

The same software was applied to study allelic and genotypic associations with the diseases, by using chi-squared tests and Cochrane Armitage analyses. The whole population of patients with RA and MG, as well as subgroups of patients harboring different autoantibodies characteristic for the given disease (anti-CCP and anti-AChR antibodies for RA and MG, respectively) were analyzed. Interactions, specifically those of age and the polymorphism, were first identified by an exploratory data analysis using Bayesian multilevel analysis (BMLA) [31]. In order to clarify this interaction and potential temporal effects of this mutation, we performed two analyses. In the first approach we investigated the association of this SNP in a stratified analysis using three age groups i.e. below 40 years, 41–69 years and above 70 years. We also examined the effect of the polymorphism adjusted for age and gender with logistic regression. Furthermore,

Table 1
Genotype and allele frequencies as well as deviation from the Hardy–Weinberg equilibrium.

Study group	Genotype 11 [a]		Genotype 12 [a]		Genotype 22 [a]		N[b]	Allele2(%)[c]	HWE[d]
Control	153	44.5%	163	47.3%	28	8.1%	344	0.32	0.1
RA	101	47.1%	79	36.9%	34	16%	214	0.39	0.07
RA 2[e]	52	38%	61	46%	21	15.7%	134		
MG	58	39%	72	49%	19	13%	149	0.37	0.64

[a] Alleles on the forward strand; 1, frequent allele (A); 2, rare allele (T) in the Caucasian population.
[b] Number of controls or patients successfully genotyped for the SNP.
[c] Frequency for allele2.
[d] HWE: indicates the significance level (p-value) of the deviation from the Hardy–Weinberg equilibrium (χ^2 test was used).
[e] Second, independent RA cohort.

in the second approach using only RA cases, we investigated the dependence of this SNP with the "age at onset" of RA. We applied ANOVA and we also used Mann–Whitney, and Kolmogorov–Smirnoff tests, because of the non-normality of the RA onset. Logistic regression and statistical tests were performed using SPSS 17.0 software (SPSS Inc.).

The analyses were performed using additive ([AA] ↔ [AT] ↔ [TT]), recessive ([AA + AT] ↔ [TT]) or dominant ([AA] ↔ [AT + TT]) models (where A was the frequent and T was the rare allele). The strength of association was represented by odds ratio (OR) with 95% confidence intervals (CI) or exponentiation of the B coefficient (Exp(B)) in logistic regression models. OR is the ratio of the odds of an event occurring in one group to the odds of it occurring in the other group. Exp(B) value is the OR in logistic regression models. (CI measures the probability ("confidence") that a population parameter will fall in a given interval in repeated studies).

3. Results

3.1. Association of rs2737713 with MG, anti-AChR antibody-positive MG, RA and anti-CCP antibody-positive RA

The overall frequencies of the presence of the SNP are presented in Table 1. Genotype distributions were in Hardy–Weinberg equilibrium for the SNP in all populations.

The patients tested in our study were grouped also according to their antibody status. Anti-cyclic citrullinated peptide antibodies (anti-CCP) are considered as highly specific biomarkers of rheumatoid arthritis present in 60–80% of patients suffering from this disease [32,33]. Autoimmune myasthenia gravis is characterized by antibodies directed against different components of the neuromuscular junction. Over 70–90% of the patients harbor anti-AChR autoantibodies representing the largest subgroup of this disease [34]. In our cohort 65% (140/214) of the patients with RA were anti-CCP antibody positive, while 70% (105/149) of the patients with MG were positive for anti-AChR antibodies. We determined the occurrence of the *rs2737713* SNP in relation to clinical parameters.

Results shown in Supplementary Table 1 reveal a strong association of RA and anti-CCP antibody-positive RA and a significant association of anti-AChR antibody-positive MG with the rs237713 polymorphism. Since age and gender distributions differed significantly between the control and the patient groups, they were controlled using logistic regression and stratification. In logistic regression models considering age and gender rs2737713 occurrence and RA ($p = 0.007$, Exp(B) coefficient (coeff.) = 4.871) as well as anti-CCP antibody-positive RA ($p = 0.008$, Exp(B) coeff. = 5.173) were found to be associated with statistical significance. In order to verify the association of rs2737713 polymorphism with RA, a separate cohort of 134 RA patients was tested. The rs2737713 TT genotype was found to be associated with the disease (28 individuals (8%) of the control group, while 21 (15.7%) RA patients harbored the TT genotype, $p = 0.04$) using logistic regression analysis (data not shown) in the independent group of patients also. In the same logistic regression set-up significance was reached for rs2737713 presence only in interaction with age for MG ($p = 0.021$, Exp(B) coeff. = 0.993) and for anti-AChR antibody positivity ($p = 0.008$, Exp(B) coeff. = 0.991).

We also performed a stratified analysis using three age groups (below 40 years, 41–69 years and above 70 years). Our analysis revealed that the rs2737713 was significantly associated with both RA and anti-CCP antibody-positive RA in the age groups below 40 and above 70 years. It also associated with MG and the anti-AChR antibody-positive subgroup of MG in the age group above 40 years (Table 2). An age dependence was noted concerning the time of disease onset and the rs2737713 polymorphism in RA patients, both in logistic regression and in stratified analyses. The common OR changed from 0.307 in the patient group under 40 years of age to 2.934 in patients over 70 years (not shown). The lack of association in the strata (40–70 years) within the RA group may be attributed to the transitional phase (Table 2). For gender correction, a stratified analysis was carried out in the women subgroup. Owing to the small sample size, men were not analyzed separately. The statistical pairing came up with significance when analyzing only women with RA (Fig. 1). Estimated RA prevalence according to age at onset in the entire RA group (Fig. 1A) and the subgroup of female patients (Fig. 1B) is shown.

Table 2
Association of the studied polymorphism with RA and MG in different age groups.

Disease state	Groups of patients according to rs2737713 compared to controls[a]	Comparison of frequencies of patients with controls		
		p value	OR	95% CI
RA < 40 years	[1] ↔ [2]	**0.028**	0.35	0.13–0.928
	[11] ↔ [22]	0.126	0.196	0.01–3.481
	[11] ↔ [12 + 22]	**0.021**	0.295	0.09–0.879
	[11 + 12] ↔ [22]	0.223	3.155	0.18–54.74
RA 40–70 years	[1] ↔ [2]	0.197	1.276	0.88–1.85
	[11] ↔ [22]	0.107	1.974	0.85–4.56
	[11] ↔ [12 + 22]	0.554	1.158	0.71–1.88
	[11 + 12] ↔ [22]	0.091	0.505	0.23–1.13
RA > 70 years	[1] ↔ [2]	**0.0069**	3.22	1.35–7.69
	[11] ↔ [22]	**0.0163**	10.28	1.15–91.62
	[11] ↔ [12 + 22]	0.0557	2.86	0.96–8.49
	[11 + 12] ↔ [22]	**0.0267**	0.12	0.01–1.029
Anti-CCP + RA < 40 years	[1] ↔ [2]	**0.047**	0.349	0.11–1.031
	[11] ↔ [22]	0.171	0.243	0.01–4.367
	[11] ↔ [12 + 22]	**0.0372**	0.295	0.088–0.983
	[11 + 12] ↔ [22]	0.274	2.55	0.14–44.732
Anti-CCP + RA 40–70 years	[1] ↔ [2]	0.495	1.14	0.78–1.66
	[11] ↔ [22]	0.559	1.264	0.58–2.77
	[11] ↔ [12 + 22]	0.558	1.161	0.71–1.91
	[11 + 12] ↔ [22]	0.637	0.84	0.39–1.76
Anti-CCP + RA > 70 years	[1] ↔ [2]	**0.0008**	3.875	1.7–8.72
	[11] ↔ [22]	**0.0009**	21	2.3–191
	[11] ↔ [12 + 22]	**0.0685**	2.69	0.91–7.9
	[11 + 12] ↔ [22]	**0.0006**	0.051	0.006–0.435
MG > 40 years	[1] ↔ [2]	**0.0092**	1.655	1.13–2.42
	[11] ↔ [22]	**0.013**	2.859	1.22–6.69
	[11] ↔ [12 + 22]	**0.033**	1.741	1.041–2.912
	[11 + 12] ↔ [22]	**0.016**	2.082	1.14–3.802
Anti-AChR + MG > 40 years	[1] ↔ [2]	**0.00435**	1.863	1.21–2.87
	[11] ↔ [22]	**0.00693**	3.5	1.37–9.00
	[11] ↔ [12 + 22]	**0.016**	2.082	1.14–3.8
	[11 + 12] ↔ [22]	**0.028**	0.385	0.16–0.925

p values < 0.05.
OR: odds ratio (measures effect size, describes the strength of association. Definition wise: the odds ratio is the ratio of the odds of an event occurring in one group to the odds of it occurring in another group. The term is also used to refer to sample-based estimates of this ratio).
CI: confidence interval: measures the probability that a population parameter will fall between two set values. $p < 0.05$ was used in this study.
[a] Alleles on the forward strand; 1, frequent allele (A); 2, rare allele (T) (in the Caucasian population).

3.2. Association of rs2737713 with RA onset

The obtained data led us to perform a related quantitative trait-association analysis for RA onset. Accordingly, rs2737713 occurrence is associated with RA onset, using the Mann–Whitney test ($p = 0.02$), the Kolmogorov–Smirnoff test ($p = 0.009$) and ANOVA ($p = 0.032$, for descriptives, please see Supplementary Table 2). Robust and efficient nonparametric tests were used because RA-onset did not follow normal distribution. Moreover, we added allelic and genotypic association tests for an indicator variable "onset before 50 years" derived from the disease onset of RA. This data processing provided evidence that the presence of the allelic TT variant significantly reduces the odds ratio of early onset RA (e.g. common odds ratio is 0.426 with $p = 0.0008$;

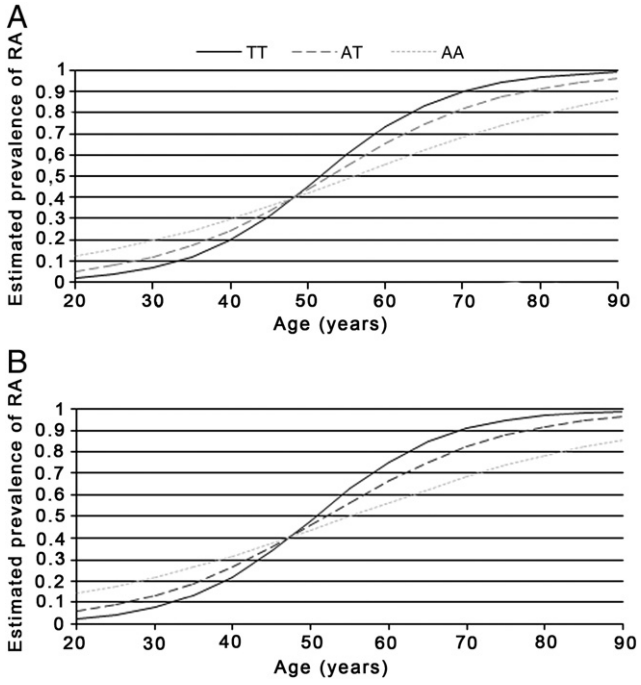

Fig. 1. Age dependence of the prevalence for RA using logistic regression for the entire investigated RA population (A) and the female RA patients (B). The diagram shows the dependence of the prevalence of the OR defined by the logistic regression model (y-axis) from age (x-axis) in both the investigated RA population (A) and in women suffering from RA according the genotype (TT, AT, AA respectively) (B). As shown, the TT genotype increases susceptibility in late-onset RA. Owing to the small sample size of male RA patients this analysis was not done in this subgroup.

for details, please see Table 3). This information prompted us to proceed to monitor features of the protein.

3.3. Association of rs2737713 TT polymorphism with protein level and lectin activity

We first measured the level of galectin-8 in PBMC lysates of genotyped healthy individuals ($n = 9$, $n = 16$ for TT/AA genotypes, respectively) in ELISA-based assays. Genotype difference was not associated with an alteration in protein presence ($p = 0.657$). Normalized signal intensities in solid-phase binding assays with lactose and 3′-sialyllactose to probe for an effect of the F19Y substitution on ligand binding showed rather similar values in comparison ($p = 0.976$, $p = 0.546$ for the two ligands; $p = 0.293$ considering the ratio; Supplementary Fig. 1a–c). These binding assays were repeated with recombinant proteins (N-terminal domain and full-length protein) using antibody/streptavidin-based detection systems. Binding was saturable and completely blocked by cognate sugar inhibitors; specificity controls with non-cognate α2,6-sialylated lactose and mannose resulted in expectedly negligible signals defining the background level, in contrast to testing reactive glycans. The antibody-based detection system revealed no difference in binding of the N-terminal domains to lactose but an increase for binding to 3′-sialyllactose (Supplementary Fig. 2a and b). Using biotinylated proteins, no firm evidence for an alteration was obtained (not shown). Testing four ligands in this system in parallel, reactivity was relatively high for 3′-sialyllactose for the N-terminal domain (Supplementary Fig. 2). Using the dimeric building block of physiological poly(N-acetyllactosamine) repeats and 3′-sulfated TF-antigen as ligands with the N-terminal domain and the full-length protein, we obtained slightly lower signals but no significant difference between wild-type and mutant proteins could be found (not shown).

Table 3
Association of rs2737713 with early onset RA.

Groups of patients according to rs2737713[a] with early disease onset compared to age matched controls[b]	Comparison of frequencies of patients with controls		
	p value	OR	95% CI
[1] ↔ [2]	**0.0003**	0.386	0.22–0.652
[11] ↔ [22]	**0.0024**	0.166	0.04–0.592
[11] ↔ [12 + 22]	**0.0027**	0.373	0.19–0.718
[11 + 12] ↔ [22]	**0.01**	4.5	1.3–15.6

p values < 0.05.
[a] Alleles on the forward strand; 1, frequent allele (A); 2, rare allele (T) (in the Caucasian population).
[b] Early onset RA < 50 years.

4. Discussion

This study was designed to put the hypothesis to the test, whether a gene for an immunoregulatory galectin may harbor a SNP in the coding region associated with autoimmune diseases. It revealed associations of the rs2737713 *LGALS8* polymorphism with both the clinical and the corresponding antibody status in patients with RA and MG. Moreover, data processing delineated that the TT genotype in the case of RA is less frequent in the younger onset (<50 years) population than in patients with older age at disease onset. Looked at in clinical terms, a genotype which appears to be beneficial at a younger age may be unfavorable later on, a phenomenon termed as antagonistic pleiotropy. Polymorphisms of genes encoding pro-inflammatory proteins have been described with such characteristics [35,36]. The complete spectrum of immunoregulatory activities of galectin-8, which is also regulated in tumor progression [37–39], has not yet been defined; the detection of this association gives further research a promising perspective.

It constitutes the first example of a SNP in a coding sequence for a galectin associated with an autoimmune disease. Thus, it moves beyond a previous report in two regulatory-region polymorphisms of galectin-1 and MG [40]. Testing of cell lysates and recombinant proteins did not reveal a marked influence of the F19Y substitution on binding certain carbohydrate ligands. The results of this study are summarized in Table 4. However, the examples of the galectin-3 rs4644 SNP (P64H), affecting proteolytic processing of this family member [41,42], which shares strong reactivity to histo-blood group A/B epitopes with galectin-8 [43], and of the C2S substitution in galectin-1 with structural long-range effects [44] as well as the reactivity of galectins with other proteins [11,12] teach the lesson that the substitution may have a bearing on activities different from binding the tested carbohydrate ligands. Fittingly, the L11A mutation in galectin-1 markedly impaired placement of oncogenic H-ras in microdomains, while binding glycan to the carbohydrate recognition

Table 4
Summary of the results.

Investigated parameter of the rs2737713 SNP		Results
Association with diseases	Rheumatoid arthritis	❖ strong association ❖ also with the anti-CCP positive subgroup ❖ frequency in distribution of genotypes in anti-CCP positive subgroup and entire RA group does not differ significantly ❖ association with age at disease onset: the rare homozygous TT genotype is more frequent in late-onset RA: a case of antagonistic pleiotropy
	Myasthenia gravis	❖ mild but significant association ❖ frequency in distribution of genotypes in anti-AChR positive subgroup and entire MG group does not differ significantly
Property of the encoded protein		❖ expression not affected ❖ binding of tested carbohydrates not altered significantly

domains remained unaffected [44], and galectin-8 is also known to be present in nuclei and cytoplasm [37,39,45,46]. In the same sense, the identification of the clinical association between the SNP and two autoimmune diseases establish the mutant protein as a target for comparative structural analysis and as an indicator for broadening our view on galectin-8 functionality. In the context of the network of galectins in immunoregulation, our study encourages screening of the coding regions of other galectins for polymorphisms, associated with autoimmune disease, hereby opening a promising research area.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.bbagen.2012.05.015.

Acknowledgements

We are grateful to all patients and volunteers of the control group, whose participation made this study possible. This study was financially supported generously by grants from the Hungarian Scientific Research Fund (OTKA 73247, OTKA 84043, PD-76348, TECH_08-A1/2-2008-0120NKTH (Genagrid), TÁMOP-4.2.1/B-09/1/KMR-2010-0002, and TÁMOP 4.2.1./B-09/1/KMR), a Marie Curie RTN (#19561), the EC GlycoHIT consortium (contract number 26060) and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences (P.A.). The sponsors did not participate in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

References

[1] A. Zhernakova, C.C. van Diemen, C. Wijmenga, Detecting shared pathogenesis from the shared genetics of immune-related diseases, Nat. Rev. Gen. 10 (2009) 43–55.
[2] I.R. MacKay, Clustering and commonalities among autoimmune diseases, J. Autoimmun. 33 (2009) 170–177.
[3] G.S. Firestein, Evolving concepts of rheumatoid arthritis, Nature 423 (2003) 356–361.
[4] A.K. Lefvert, Non-MHC genetic polymorphisms with functional importance for human myasthenia gravis, in: M. Zouali (Ed.), Molecular Autoimmunity, Springer US, New York, 2005, pp. 101–113.
[5] G. Orozco, B. Rueda, J. Martin, Genetic basis of rheumatoid arthritis, Biomed. Pharmacother. 60 (2006) 656–662.
[6] International MHC and Autoimmunity Genetics Network, J.D. Rioux, P. Goyette, T.J. Vyse, L. Hammarström, M.M. Fernando, et al., Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases, Proc. Natl. Acad. Sci. U.S.A. 106 (2009) 18680–18685.
[7] Wellcome Trust Case Control Consortium, Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls, Nature 447 (2007) 661–678.
[8] H.-J. Gabius, Cell surface glycans: the why and how of their functionality as biochemical signals in lectin-mediated information transfer, Crit. Rev. Immunol. 26 (2006) 43–80.
[9] R. Schwartz-Albiez, Inflammation and glycosciences, in: H.-J. Gabius (Ed.), The Sugar Code. Fundamentals of Glycosciences, Wiley VCH, Weinheim, Germany, 2009, pp. 447–467.
[10] H.-J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, D. Solís, From lectin structure to functional glycomics: principles of the sugar code, Trends Biochem. Sci. 36 (2011) 298–313.
[11] A. Villalobo, A. Nogales-Gonzáles, H.-J. Gabius, A guide to signaling pathways connecting protein–glycan interaction with the emerging versatile effector functionality of mammalian lectins, Trends Glycosci. Glycotechnol. 18 (2006) 1–37.
[12] H. Kaltner, H.-J. Gabius, A toolbox of lectins for translating the sugar code: the galectin network in the phylogenesis and tumors, Histol. Histopathol. 27 (2012) 397–416.
[13] Y. Zick, M. Eisenstein, R.A. Goren, Y.R. Hadari, Y. Levy, D. Ronen, Role of galectin-8 as a modulator of cell adhesion and cell growth, Glycoconj. J. 19 (2004) 517–526.
[14] L. Eshkar Sebban, D. Ronen, D. Levartovsky, O. Elkayam, D. Caspi, S. Aamar, H. Amital, A. Rubinow, I. Golan, D. Naor, Y. Zick, I. Golan, The involvement of CD44 and its novel ligand galectin-8 in apoptotic regulation of autoimmune inflammation, J. Immunol. 179 (2007) 1225–1235.
[15] L. Massardo, C. Metz, E. Pardo, V. Mezzano, M. Babul, E. Jarpa, A.M. Guzmán, S. André, H. Kaltner, H.-J. Gabius, S. Jacobelli, A. González, A. Soza, Autoantibodies against galectin-8: their specificity, association with lymphopenia in systemic lupus erythematosus and detection in rheumatoid arthritis and acute inflammation, Lupus 18 (2009) 539–546.
[16] H. Ideo, A. Seko, I. Ishizuka, K. Yamashita, The N-terminal carbohydrate recognition domain of galectin-8 recognizes specific glycosphingolipids with high affinity, Glycobiology 13 (2003) 713–723.
[17] S. Carlsson, C.T. Öberg, M.C. Carlsson, A. Sundin, U.J. Nilsson, D. Smith, R.D. Cummings, J. Almkvist, A. Karlsson, H. Leffler, Affinity of galectin-8 and its carbohydrate recognition domains for ligands in solution and at the cell surface, Glycobiology 17 (2007) 663–676.
[18] S.R. Stowell, C.M. Arthur, K.A. Slanina, J.R. Horton, D.F. Smith, R.D. Cummings, Dimeric galectin-8 induces phosphatidylserine exposure in leukocytes through polylactosamine recognition by the C-terminal domain, J. Biol. Chem. 283 (2008) 20547–20559.
[19] H. Kaltner, D. Solís, S. André, M. Lensch, J.C. Manning, M. Mürnseer, J.L. Sáiz, H.-J. Gabius, Unique chicken tandem-repeat-type galectin: implications of alternative splicing and a distinct expression profile compared to those of the three proto-type proteins, Biochemistry 48 (2009) 4403–4416.
[20] D. Solís, M.J. Maté, M. Lohr, J.P. Ribeiro, L. López-Merino, S. André, E. Buzamet, F.J. Cañada, H. Kaltner, M. Lensch, F.M. Ruiz, G. Haroske, U. Wollina, M. Kloor, J. Kopitz, J.L. Sáiz, M. Menéndez, J. Jiménez-Barbero, A. Romero, H.-J. Gabius, N-domain of human adhesion/growth regulatory galectin-9: preference for distinct conformers and non-sialylated N-glycans and detection of ligand-induced structural changes in crystal and solution, Int. J. Biochem. Cell Biol. 42 (2010) 1019–1029.
[21] C.-M. Tsai, C.-H. Guan, H.-W. Hsieh, T.-L. Hsu, Z. Tu, K.-J. Wu, C.-H. Lin, K.-I. Lin, Galectin-1 and galectin-8 have redundant roles in promoting plasma cell formation, J. Immunol. 187 (2011) 1643–1652.
[22] O.A. Vokhmyanina, E.M. Rapoport, I.M. Ryzhov, E.Y. Korchagina, G.V. Pazynina, V.V. Severov, H. Kaltner, S. André, H.-J. Gabius, N.V. Bovin, Carbohydrate specificity of chicken and human tandem-repeat-type galectins-8 in composition of cells, Biochemistry (Mosc) 76 (2011) 1185–1192.
[23] M.V. Tribulatti, V. Cattaneo, U. Hellman, J. Mucci, O. Campetella, Galectin-8 provides costimulatory and proliferative signals to T lymphocytes, J. Leukoc. Biol. 86 (2009) 371–380.
[24] V. Cattaneo, M.V. Tribulatti, O. Campetella, Galectin-8 tandem-repeat structure is essential for T-cell proliferation but not for co-stimulation, Biochem. J. 434 (2011) 153–160.
[25] M.F. López-Lucendo, D. Solís, S. André, J. Hirabayashi, K.-I. Kasai, H. Kaltner, H.-J. Gabius, A. Romero, Growth regulatory human galectin-1: crystallographic characterization of the structural changes induced by single-site mutations and their impact on the thermodynamics of ligand binding, J. Mol. Biol. 343 (2004) 957–970.
[26] M.J. Molnar, P. Bencsik, Establishing a neurological–psychiatric biobank: banking, informatics, and ethics, Cell. Immunol. 244 (2006) 101–104.
[27] D. Aletaha, T. Neogi, A.J. Silman, J. Funovits, D.T. Felson, C.O. Bingham CO III, et al., Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative, Arthritis Rheum. 62 (2010) 2569–2581.
[28] A. Beer, S. André, H. Kaltner, M. Lensch, S. Franz, K. Sarter, C. Schulze, U.S. Gaipl, P. Kern, M. Hermann, H.-J. Gabius, Human galectins as sensors for apoptosis/necrosis-associated surface changes of granulocytes and lymphocytes, Cytometry A 73 (2008) 139–147.
[29] S. André, Z. Pei, H.-C. Siebert, O. Ramström, H.-J. Gabius, Glycosyldisulfides from dynamic combinatorial libraries as O-glycoside mimetics for plant and mammalian lectins: their reactivities in solid-phase and cell assays and conformational analysis by molecular dynamics simulations, Bioorg. Med. Chem. 14 (2006) 6314–6323.
[30] J.E. Wigginton, D.J. Cutler, G.R. Abecasis, A note on exact tests of Hardy–Weinberg equilibrium, Am. J. Hum. Genet. 76 (2005) 887–893.
[31] P. Antal, A. Millinghoffer, G. Hullám, C. Szalai, A. Falus, A Bayesian view of challenges in feature selection: feature aggregation, multiple targets, redundancy and interaction. ECML/PKDD workshop on new challenges for feature selection in data mining and knowledge discovery (FSDM08), JMLR: Workshop and Conference Proceedings, 4, 2008, pp. 74–89.
[32] N. Inanc, E. Dalkilic, S. Kamali, E. Kasapoglu-Günal, Y. Elbir, H. Direskeneli, M. Inanc, Anti-CCP antibodies in rheumatic arthritis and psoriatic arthritis, Clin. Rheumatol. 1 (2007) 17–23.
[33] M. Puszczewicz, C. Iwaszkiewicz, Role of anti-citrullinated protein antibodies in diagnosis and prognosis of rheumatoid arthritis, Arch. Med. Sci. 2 (2011) 189–194.
[34] A. Vincent, J. McConville, M.E. Farrugia, J. Bowen, P. Plested, T. Tang, A. Evoli, I. Matthews, G. Sims, P. Dalton, L. Jacobson, A. Polizzi, F. Blaes, B. Lang, D. Beeson, N. Willcox, J. Newsom-Davis, W. Hoch, Antibodies in myasthenia gravis and related disorders, Ann. N. Y. Acad. Sci. 998 (2003) 324–335.
[35] S. Vasto, G. Candore, C.R. Balistreri, M. Caruso, G. Colonna-Romano, M.P. Grimaldi, F. Listi, D. Nuzzo, D. Lio, C. Caruso, Inflammatory networks in ageing, age-related diseases and longevity, Mech. Ageing Dev. 128 (2007) 83–91.
[36] C. Franceschi, M. Bonafè, S. Valensin, F. Olivieri, M. De Luca, E. Ottaviani, G. De Benedictis, Inflamm-aging. An evolutionary perspective on immunosenescence, Ann. N. Y. Acad. Sci. 908 (2000) 244–254.
[37] N. Nagy, Y. Bronckart, I. Camby, H. Legendre, H. Lahm, H. Kaltner, Y. Hadari, P. Van Ham, P. Yeaton, J.-C. Pector, Y. Zick, I. Salmon, A. Danguy, R. Kiss, H.-J. Gabius, Galectin-8 expression decreases in cancer compared with normal and dysplastic human colon tissue and acts significantly on human colon cancer cell migration as a suppressor, Gut 50 (2002) 392–401.
[38] S. Langbein, J. Brade, J.K. Badawi, M. Hatzinger, H. Kaltner, M. Lensch, K. Specht, S. André, U. Brinck, P. Alken, H.-J. Gabius, Gene-expression signature of adhesion-growth regulatory tissue lectins (galectins) in transitional cell cancer and its prognostic relevance, Histopathology 51 (2007) 681–690.
[39] S. Cludts, C. Decaestecker, V. Mahillon, D. Chevalier, H. Kaltner, S. André, M. Remmelink, X. Leroy, H.-J. Gabius, S. Saussez, Galectin-8 upregulation during hypopharyngeal and laryngeal tumor progression and comparison with galectins-1, -3 and -7, Anticancer. Res. 29 (2009) 4933–4940.

[40] Z. Pal, P. Antal, A. Millinghoffer, G. Hullám, K. Pálóczi, S. Tóth, H.-J. Gabius, M.J. Molnar, A. Falus, E.I. Buzás, A novel galectin-1 and interleukin-2 receptor β haplotype is associated with autoimmune myasthenia gravis, J. Neuroimmunol. 229 (2010) 107–111.
[41] M.M. Lotz, C.W. Andrews Jr., C.A. Korzelius, E.C. Lee, G.D. Steele Jr., A. Clarke, A.M. Mercurio, Decreased expression of Mac-2 (carbohydrate binding protein 35) and loss of its nuclear localization are associated with the neoplastic progression of colon carcinoma, Proc. Natl. Acad. Sci. USA 90 (1993) 3466–3470.
[42] V. Balan, P. Nangia-Makker, A.G. Schwartz, Y.S. Jung, L. Tait, V. Hogan, T. Raz, Y. Wang, Z.Q. Yang, G.S. Wu, Y. Guo, H. Li, J. Abrams, F.J. Couch, W.L. Lingle, R.V. Lloyd, S.P. Ethier, M.A. Tainsky, A. Raz, Racial disparity in breast cancer and functional germ line mutation in galectin-3 (rs4644): a pilot study, Cancer Res. 68 (2008) 10045–10050.
[43] M. Krzeminski, T. Singh, S. André, M. Lensch, A.M. Wu, A.M.J.J. Bonvin, H.-J. Gabius, Human galectin-3 (Mac-2 antigen): defining molecular switches of affinity to natural glycoproteins, structural and dynamic aspects of glycan binding by flexible ligand docking and putative regulatory sequences in the proximal promoter region, Biochim. Biophys. Acta 1810 (2011) 150–161.
[44] B. Rotblat, H. Niv, S. André, H. Kaltner, H.-J. Gabius, Y. Kloog, Galectin-1 (L11A) predicted from a computed galectin-1 farnesyl-binding pocket selectively inhibits Ras-GTP, Cancer Res. 64 (2004) 3112–3118.
[45] S. Saussez, L. de Leval, C. Decaestecker, N. Sirtaine, S. Cludts, A. Duray, D. Chevalier, S. André, H.-J. Gabius, M. Remmelink, X. Leroy, Galectin fingerprinting in Warthin's tumors: lectin-based approach to trace its origin? Histol. Histopathol. 25 (2010) 541–550.
[46] M. Remmelink, L. de Leval, C. Decaestecker, A. Duray, E. Crompot, N. Sirtaine, S. André, H. Kaltner, X. Leroy, H.-J. Gabius, S. Saussez, Quantitative immunohistochemical fingerprinting of adhesion/growth-regulatory galectins in salivary gland tumors: divergent profiles with diagnostic potential, Histopathology 58 (2011) 543–556.