Evaluation of Galectin-3 Genetic Variants and its Serum Levels in Rheumatoid Arthritis in North India

Tarnjeet Kaur1, Annie Sodhi1, Jatinder Singh2, Sumeet Arora3, Sukhdev Singh Kamboj2 and Manpreet Kaur1*

1Department of Human Genetics, Guru Nanak Dev University, Amritsar143005, Punjab, India
2Department of Molecular Biology and Biochemistry, Guru Nanak Dev University, Amritsar143005, Punjab, India
3Medical and Arthritis Care Centre, Amritsar143001, Punjab, India

KEYWORDS Atherogenic Index. C-Reactive Protein. Inflammation. Lipid Profile. Single Nucleotide Polymorphism

ABSTRACT The aim of present study was to evaluate the role of galectin-3 variants (rs4644 and rs4652) and its serum levels as predisposition factor for Rheumatoid Arthritis (RA) in North India. The present study recruited 200 RA patients and 200 unrelated age, gender and ethnicity matched controls. Genotyping of rs4644 (+191C/A) and rs4652 (+292A/C) was performed by PCR- RFLP. Serum galectin-3 levels and hs-CRP (high sensitivity-C reactive protein) levels were assessed using ELISA kits. Different lipid profile biomarkers were quantified using standard reagents and kits. LGALS3 +191(C>A) showed significant difference (p<0.05) in genotypic distribution between patients and controls. Patients were found to have high serum galectin-3 levels, high atherogenic index and higher levels of hs-CRP (p<0.01). These results indicate that genetic polymorphism in galectin-3 gene may contribute to development of RA.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic, inflammatory, systemic disease characterized by destruction of peripheral joints leading to deformity and disability. About 1 percent of the human population worldwide is suffering from this disease (Feldmann et al. 1996). Dysregulation of CD4+ T-cells is proposed to be one of the key event in pathogenesis of RA, which can amplify the immune response by stimulating other mononuclear cells, synovial fibroblasts, chondrocytes and osteoclasts (Firestein 2004; Mclnnes and Schett 2011). A number of inflammatory mediators have been implicated in the establishment and progression of inflammatory joint destruction, including various pro-inflammatory cytokines like interleukin-1(IL-1), interleukin-6 (IL-6) and tumour necrosis factor-α (TNF-α) as well as proteins from a novel super-family of animal lectins, the galectins (Almkvist and Karlsson 2004; Leffler et al. 2004). Galectin-3 is a 32kDa chimera-type member of the galectin family and is encoded by LGALS3 gene located on chromosome 14q21. It consists of a CRD and an N-terminal domain that are responsible for interactions between subunits facilitating its oligomerization (Dumic et al. 2006). This oligomerized galectin-3 binds with high affinity to terminal N-acetyllactosamine moieties on T-cell receptor (TCR). This interaction of galectin-3 with these moieties was proposed to regulate T-cell activation by limiting TCR clustering necessary for TCR mediated signal initiation (Hsu et al. 2009). Furthermore, galectin-3 selectively promoted the secretion of mononuclear cell–specific chemokines from synovium, but not skin fibroblasts in RA, suggesting that galectin-3 can amplify joint inflammation (Filer et al. 2009). Also high expression of...
galectin-3 and its binding protein in RA synovi-um demonstrated its plausible role in pathogen-esis of RA (Ohshima et al. 2003; Forsman et al. 2011).

Objectives

The objective of the present study was to test whether the genetic polymorphisms of ga-lectin-3 confers predisposition to RA. To meet this objective, galectin-3 genetic variants and its serum levels were investigated and compared between RA patients and healthy subjects from North India. To the best of the researchers’ knowledge it is the first report of galectin-3 gene polymorphism in RA in India.

METHODOLOGY

The present case control association study enrolled 200 RA patients diagnosed according to 1987 revised criteria of American college of Rheumatology (Arnett et al. 1988). All the patients were recruited from a local rheumatology clinic and information regarding demographic features including age, gender, occurrence of hypertension or diabetes mellitus, disease dura-tion, erythrocyte sedimentation rate (ESR), rheu-matoid factor (RF) etc. was also collected from patients. A cohort of 200 age and gender matched (p>0.05) healthy individuals, having no known medical history on a health screening questionnaire, were recruited as controls. All the participants were genetically unrelated North Indian subjects. Pregnant women, diabetics, alcoholics, smokers as well as those having known cardiovascular disease and primary muscle dis-ease were excluded from the study. The study was approved by institutional ethical committee in accordance with declaration of Helsinki and written informed consent was obtained from each individual. Five ml of blood was withdrawn from all the subjects and it was processed for DNA and serum isolation. The genomic DNA was iso-lated from peripheral blood mononuclear cells (PBMC) by inorganic method and stored at -80°C till further use (Miller et al. 1988). The nucle-o-tide sequence harboring rs4644(+191 C/A) and rs4652(+292A/C) single nucleotide polymor-phisms (SNPs) in exon3 of LGALS3 gene was amplified by polymerase chain reaction (PCR) in thermocycler (Eppendorf, Germany). The primers used for PCR were 5’GTATGCTTTTCTT- CAGCTC3’ (forward) and 5’ CAGTTTTAT- CAGTGCCCTC 3’ (reverse) (Primer- BLAST).

PCR was performed in a final volume of 20 μl using standardized protocol containing 50 ng DNA, 1X Taq buffer A with 15 mM MgCl₂, 200 μM dNTPs, 0.3 μM primer mix (forward and reverse primer each), and 0.3 U Taq polymerase. The PCR protocol consists of 3 min initial dena-turation at 94°C and then 35 cycles each of 30 s denaturation at 94°C, 30 s annealing at 60.2°C, 30 s extension at 72°C and single step of final extension at 72°C for 7 min. Genotypes were determined by restriction fragment length polymorphism (RFLP). The 553 bp PCR product obtained was further digested with 1U NcoI restriction enzyme (New England Biolabs, USA) for rs4644 SNP and with 1U BsaW1 (New En-gland Biolabs, USA) for rs4652 at 37°C for 16 hours. The digested products obtained were analyzed on 3% agarose gel. For rs4644 SNP, the products of sizes 443, 110 bp for CC genotype; 443, 254, 189, 110 bp for CA genotype and 254, 189, 110 bp were obtained for AA genotype (Fig. 1). In case of rs4652 SNP 292, 261 bp products for AA genotype; 553, 292, 261 bp products for AC genotype and 553 bp product for CC geno-type were obtained (Fig. 1).

Serum galectin-3 levels (n=110) were as-sessed in cases as well as controls using com-mercially available enzyme linked immune-sor-bent assay (ELISA) kits (Abcam, UK). Briefly, 100μl of pre-diluted sample was added to mi-crotitre plate pre-coated with anti-human galec-tin-3 antibody followed by addition of biotin-conjugated anti-human galectin-3 antibody and the plate was incubated at 37°C for 2 hours. Following incubation, unbound biotin conjugated anti-humanantibodywas removed by washing with 1X wash buffer and streptavidin-HRP was added to all wells followed by incubation at 37°C for 2 hours. After 3-4 washings, substrate was added and the coloured complex obtained was measured at 450nm. High sensitivity-C-reactive protein (hs-CRP) levels (n=80) were also es-timated in serum samples of cases and controls using ELISA kits (MyBioSource, USA).

Serum lipid profile biomarkers i.e. choles-terol, triglycerides and high density lipoprotein (HDL) were analyzed on semiautomatic clinical analyzer (ERBA, Germany) while apo-lipopro-tein A1 (apoA1) and apo-lipoprotein B (apoB)
were analyzed on spectrophotometer (Eppendorf, Germany) using standard reagents and kits (ERBA, Germany). Very low density lipoprotein (VLDL) and low density lipoprotein (LDL) were calculated from the above mentioned parameters by standard formulae (Friedewald et al. 1972).

Sample size for genetic association was calculated by CaTS Power calculator to achieve a power of 80 percent, taking assumptions of 1 percent worldwide prevalence and an odds ratio of 1.5 ($\alpha$ = 0.05). Genotypic and allelic frequencies were compared between RA patients and controls using odds ratio. Analysis was also performed assuming dominant, co-dominant and recessive genetic models using Web-Assotest program. Different genetic combinations for both SNPs were estimated in cases and controls by PHASE software version 2.1. Shapiro-Wilk test was used to check normality of data. Serum galectin-3, hs-CRP and lipid profile variables were compared between cases and controls using Mann Whitney U test. Comparison of galectin-3 levels in different genetic combinations of cases and controls was done by two-way ANOVA. All outliers and extreme values from whole data were detected by box whisker plot and then values were adjusted according to respective means. All statistical analysis was performed using Statistical Package for the Social Sciences (SPSS) version 18.0 (IL, USA, and Chicago). p<0.05 values were considered significant.

Multiple sequence alignment for the evolutionary history of LGALS3 (+191 C/A) was performed by SWISS-MODEL. The secondary structure of galectin-3 protein was determined by PSIPRED. The different structural and functional annotations of galectin-3 protein were predicted at PREDICT PROTEIN server. The change in mRNA structure was determined at Genebee website.

RESULTS

The present study included 185 females and 15 males (mean age 47.3±12.3 years). All patients were RF positive and normo-tensive. The mean ESR of patients was 50.23±27.5 mm/h and average disease duration of patients was 5.23±3.5 year. The control group also included 185 females and 15 males (mean age 47.44±12.4 years). Serum galectin-3 levels were found to be significantly high ($p<0.001$) in cases as compared to controls. Serum hs-CRP levels were also found to be significantly higher ($p<0.001$) in patients than controls. Furthermore, the patients were found to be dyslipidemic, as indicated by significantly elevated levels of cholesterol, LDL cholesterol, TG/HDL ratio, LDL/HDL ratio, apoA (p
LGALS3 (+191C/A) showed significant difference (OR-0.55, 95%CI-0.27-0.98, p-0.04) in genotypic distribution between cases and controls (Table 2). The frequency of homozygous variant (AA) was significantly higher in cases (25.5%) as compared to controls (20%) while frequency of CC genotype was significantly high-

<table>
<thead>
<tr>
<th>Variables</th>
<th>RA patients (Mean±SD)(n=200)</th>
<th>Controls (Mean±SD)(n= 200)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galectin-3 (ng/ml)</td>
<td>1.94± 1.1</td>
<td>1.04± 0.7</td>
<td>0.001**</td>
</tr>
<tr>
<td>hs-CRP (mg/L)</td>
<td>8.55± 3.2</td>
<td>6.33± 1.3</td>
<td>0.001*</td>
</tr>
<tr>
<td>Lipid Profile Variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholesterol(mg/dl)</td>
<td>223.62± 51.1</td>
<td>198.48± 45.1</td>
<td>0.001**</td>
</tr>
<tr>
<td>Triglycerides(mg/dl)</td>
<td>201.65± 79.2</td>
<td>191.12± 83.5</td>
<td>0.177</td>
</tr>
<tr>
<td>HDL-cholesterol(mg/dl)</td>
<td>34.83± 17.4</td>
<td>45.76± 17.4</td>
<td>0.001**</td>
</tr>
<tr>
<td>LDL(mg/dl)</td>
<td>143.47± 52.1</td>
<td>113.91± 46.3</td>
<td>0.001**</td>
</tr>
<tr>
<td>VLDL(mg/dl)</td>
<td>40.33± 15.8</td>
<td>38.22± 16.7</td>
<td>0.177</td>
</tr>
<tr>
<td>TG/HDL</td>
<td>7.19± 4.9</td>
<td>4.71± 2.7</td>
<td>0.001**</td>
</tr>
<tr>
<td>LDL/HDL</td>
<td>5.26± 3.5</td>
<td>2.88± 1.7</td>
<td>0.001**</td>
</tr>
<tr>
<td>LDL/VLDL</td>
<td>3.89± 2.1</td>
<td>3.51± 2.2</td>
<td>0.03*</td>
</tr>
<tr>
<td>apoA1 (mg/dl)</td>
<td>72.13± 41.6</td>
<td>87.05± 54.0</td>
<td>0.009**</td>
</tr>
<tr>
<td>apoB   (mg/dl)</td>
<td>100.29± 38.5</td>
<td>92.42± 40.7</td>
<td>0.020*</td>
</tr>
</tbody>
</table>

Table 2: Distribution of genotypic and allelic frequency along with different genetic models for rs4644 and rs4652 polymorphisms of LGALS3

<table>
<thead>
<tr>
<th>Genotype/allele</th>
<th>Cases (n=200) (n (%)</th>
<th>Controls (n=200) (n (%)</th>
<th>OR, (95%CI) p-value</th>
<th>Dominant model (CA/AA vs. CC/OR (95%CI) p value</th>
<th>Co-dominant model (AA vs. CA/OR (95%CI) p value</th>
<th>Recessive model (AA vs. CC/CA/OR (95%CI) p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>28 (14)</td>
<td>42 (21)</td>
<td>0.55</td>
<td>1.63 (0.97-2.67) p value</td>
<td>1.00 (0.86-1.88) p value</td>
<td>0.80 (0.67-1.80) p value</td>
</tr>
<tr>
<td>CA</td>
<td>121 (60.5)</td>
<td>118 (59)</td>
<td>(0.27-0.98) p value</td>
<td>(0.97-2.76) p value</td>
<td>(1.00-1.88) p value</td>
<td>0.67 (0.51-1.54) p value</td>
</tr>
<tr>
<td>AA</td>
<td>51 (25.5)</td>
<td>40 (20)</td>
<td>0.04</td>
<td>0.06 (0.05-1.02) p value</td>
<td>0.05 (0.05-1.02) p value</td>
<td>0.08 (0.05-1.02) p value</td>
</tr>
</tbody>
</table>

Table 2: Distribution of genotypic and allelic frequency along with different genetic models for rs4644 and rs4652 polymorphisms of LGALS3

<table>
<thead>
<tr>
<th>Genotype/allele</th>
<th>Cases (n=200) (n (%)</th>
<th>Controls (n=200) (n (%)</th>
<th>OR, (95%CI) p-value</th>
<th>Dominant model (CA/AA vs. CC/OR (95%CI) p-value</th>
<th>Co-dominant model (AA vs. CA/OR (95%CI) p-value</th>
<th>Recessive model (AA vs. CC/CA/OR (95%CI) p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>2 (1)</td>
<td>8 (4)</td>
<td>0.27</td>
<td>4.12 (0.86-19.67) p value</td>
<td>0.89 (0.67-1.80) p value</td>
<td>0.67 (0.51-1.54) p value</td>
</tr>
<tr>
<td>AC</td>
<td>170 (85)</td>
<td>161 (80.5)</td>
<td>0.12</td>
<td>0.05</td>
<td>0.70</td>
<td>0.97</td>
</tr>
<tr>
<td>CC</td>
<td>28 (14)</td>
<td>31 (15.5)</td>
<td>0.73</td>
<td>0.33 (0.11-1.02) p value</td>
<td>0.83</td>
<td>0.19</td>
</tr>
<tr>
<td>A</td>
<td>174 (43.5)</td>
<td>177 (44.3)</td>
<td>0.83</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>
er in controls (21%) as compared to cases (14%). The allelic frequency of A allele was also found to be marginally higher in patients (55.7%) as compared to controls (49.5%). There was suggestive evidence of an association in a Co-dominant model (AA vs. CA = CA vs. CC; OR= 1.37, 95% CI 1.00-1.88, p = 0.05) (Table 2). However, the genotypic and allelic distribution of LGALS3 +291 (A>C) was similar in cases and controls (OR-0.27, 95% CI-0.05-1.41, p - 0.12). Present study also found a suggestive evidence of dominant mode of association of rs4652 with RA (AC/CC vs. AA; OR-4.12, 95% CI 0.86 -19.67, p- 0.05) (Table 2).

Haplotypic analysis showed nine different types of genetic combinations in cases and controls i.e. CAAC (48%), CCAC (17.5%), AAAC (15%), CACC (9.5%), AACCC (3.5%), CCCC (2%), AAA and CAAA (1.5% each) in controls, while in RA patients frequency of different combinations were as ; CAAC (51%), CCAC (13%), AAAC (20.5%), CACC (7%), AACCC (5.5%), CCCC (2%), AAA and CAAA (0.5% each). Heterozygosity for both SNPs (CAAC) was found to be prevalent in both cases and controls (51% in cases and 48% in controls). The frequency of rs4644 homozygous variant (AAAC, AACCC) was more prevalent in cases as compared to controls (20.5%, 5.5% respectively in cases vs. 15%, 3.5% respectively in controls). However, homozygosity for C allele (CCAC) was observed to be higher in controls than cases (17.5% vs. 13%). But the association of any of genetic combination towards risk of disease was not found to be significant (p>0.05). Further, serum galectin-3 levels stratified according to genetic combinations were compared in cases and controls. Only those combinations were taken in considerations which were represented by significant number of patients, that is, AAAC, CCAC and CAAC. For all these genetic combinations tested, cases were found to have high serum levels of galectin-3 as compared to respective controls. However, the difference was not found to be significant (p>0.05) (Table 3).

**DISCUSSION**

Galectin-3 can be viewed as regulatory switch for different biological activities including turnover of endocytic receptors, signal transduction pathways, T-cell and B-cell activation, tolerance, neutrophil activation leading to phagocytosis and oxidative burst (Kuwabara et al. 1996; Sano et al. 2000; Ochieng et al. 2004; Dumic et al. 2006). The objective of present study was to evaluate galectin-3 variants along with its serum levels in rheumatoid arthritis patients from North India.

In the present study, rs4644 SNP was found to be associated with susceptibility to RA in North India. This is indicated from significantly high (p<0.05) prevalence of AA genotype in RA patients as compared to controls. This variant lies at 64th position in N-terminal domain of galectin-3 and involves substitution of C with A, replacing proline with histidine in galectin-3 protein. His64 renders galectin-3 molecule susceptible towards cleavage by matrix metallo-proteinases (MMPs) resulting in loss of self association property and alterations in glycan binding properties (Nangia-Makker et al. 2007). So, based on this hypothesis, it has been proposed that in the absence of oligomerization, galectin-3 was unable to form galectin-glycoprotein lattices. This can subsequently lead to dys-regulation of various regulatory pathways including failure of T cell regulation or tolerance. Since T cells are one of the major cells of pathogenic importance in RA, TCR hyper-activation was proposed as one of the major pathway in pathogenesis of RA (Demetriou et al. 2001; Chen et al. 2009). Furthermore, this variation in N-terminal domain may leads to its altered secretion (Menon and Hugh-

### Table 3: Distribution of galectin-3 concentration in different genetic combinations in cases and controls

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Haplotype</th>
<th>Frequency in cases</th>
<th>Frequency in controls</th>
<th>Galectin-3 (ng/ml) Cases (Mean±SD)</th>
<th>Galectin-3 (ng/ml) Controls (Mean±SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AAAC</td>
<td>40</td>
<td>19</td>
<td>2.41± 1.25</td>
<td>1.38± 0.92</td>
<td>0.342</td>
</tr>
<tr>
<td>2</td>
<td>CAAC</td>
<td>22</td>
<td>44</td>
<td>1.47± 0.94</td>
<td>0.99± 0.54</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>CCAC</td>
<td>26</td>
<td>27</td>
<td>1.69± 0.85</td>
<td>0.93± 0.55</td>
<td></td>
</tr>
</tbody>
</table>
es 1999; Hu et al. 2011). However, for rs4652 SNP, the researchers have not found any significant difference in allelic and genotypic distribution between cases and controls. These findings are in contrast to the single report of galectin-3 polymorphism in RA in Taiwanese population, in which no association was observed with rs4644, while significant association was found for rs4652 (Hu et al. 2011). These variations can be attributed to differences in genetic constitution, ethnicity, cultures, environmental factors etc.

Comparison of +191(C/A) SNP at structural level of protein showed that molecule having Pro was marginally more stable over having His. The most plausible explanation for this might be that Pro at 64th position is evolutionarily conserved and it has been confirmed with multiple sequence alignment that any change in this position may affect the stability of this molecule. Sequence alignment also revealed that proline, glycine and alanine amino acids are abundantly present in N-terminal of galectin-3. No major changes were observed in protein functional annotations and in mRNA structure of galectin-3.

The present study also found significantly high serum levels of galectin-3 in RA patients than healthy subjects. This finding is in consonance with the previous study by Ohshima et al. (2003) indicating high serum galectin-3 levels in RA patients. In contrast to this, no difference has been observed in galectin-3 levels between patients and controls in Taiwanese population (Hu et al. 2011). Extracellular galectin-3 also been shown to have pro-inflammatory role in various inflammatory diseases (Ochieng et al. 2004). Furthermore, galectin-3 has also been shown to have inductive effect on secretion of various pro-inflammatory cytokines like IL-6 and IL-8 in rheumatoid synovial fibroblasts via selective signaling pathways (Filer et al. 2009). Also down-regulation of galectin-3 resulted in inhibition of pro-inflammatory cytokines by human monocyte derived dendritic cells (Chen et al. 2015). Galectin-3 also induces the production of reactive oxygen species (ROS) in human neutrophils and promotes chemotaxis in monocytes (Sano et al. 2000; Nieminen et al. 2005). ROS production along with lipid peroxidation products and oxidized low density lipoproteins (LDL) was shown to result in cartilage matrix degradation, inducing expression of matrix degrading enzymes and ultimately form a vicious cycle to exaggerate the inflammatory processes (Henrotin et al. 2003; Henrotin et al. 2005). So, higher levels of galectin-3 in the present study may implicate its role in pathogenesis of RA. Evidence for this also comes from the different studies showing that galectin-3 has been involved in various key processes in pathogenesis of RA including angiogenesis of endothelial cells, pannus formation, secretion of chemokines and cytokines etc. (Nangia-Makker et al. 2000; Filer et al. 2009).

Out of nine different genetic combinations observed, the frequency of AAAC genetic combination was found to be higher in patients (20.5%) than controls (15%). Furthermore, this genetic combination was found to have highest galectin-3 levels among the different combinations tested. This finding may propose the role of AAAC genetic combination as predisposition factor towards susceptibility of disease.

Severe dyslipidemia, indicated by high atherogenic index was also observed in RA patients. Several earlier studies are in agreement with this finding while contrasting reports are also documented in literature (Georgiadis et al. 2006; Kowsalya et al. 2011; Vinapamula et al. 2013). Variations in these findings might be attributed to kind of medication patients were taking at the time of study or duration of medication. The patients in this study were not taking corticosteroids or statins for the last six months of sample collection. Studies have also reported that hypercholesterolemia and elevated LDL-cholesterol levels are positively correlated with development of atherosclerosis, while elevated HDL has its athero-protective function (Ledwozym et al. 1986). So elevated cholesterol, LDL cholesterol and reduced HDL levels in the patients in this study indicated the risk of developing cardiovascular diseases (CVD) in these RA patients. Furthermore, dyslipidemia have been shown to exaggravate inflammatory process (Drechsler et al. 2010). This observation was also supported by the finding in the present study indicating that RA patients were found to have significantly high levels of serum hs-CRP, a well-established inflammatory marker and predictor of coronary disease (Ross 1999; Hansson 2001).

CONCLUSION

The present study proposes that LGALS3 +191(C/A) SNP may be associated with suscep-
tibility towards RA. The results of present study indicated the higher frequency of homozygous variant of rs4644 SNP of galectin-3 along with its higher serum galectin-3 levels in RA patients. Further, high atherogenic index and hs-CRP levels indicated more risk of CVD in the RA patients in this study.

RECOMMENDATIONS

This study involves evaluation of SNPs in N-terminal domain of galectin-3 gene. To further explore and validate the role of galectin-3 genetic polymorphism in pathogenesis of RA, further studies targeting promoter region, CRD domain as well as 3’ UTR are required.

LIMITATIONS

The present study enrolled a relatively small sample size and North Indian population. So, the conclusion of this study warrants further confirmation with comparatively larger sample size and other population groups.

ACKNOWLEDGEMENTS

This work was supported by the grant of Department of Science and Technology (DST) [SB/FT/LS-184/2012] and University Grants Commission (UGC) [F.17-74/2008(SA1)].

REFERENCES


