

Bhawna Gupta · Charu Agrawal · Sunil K. Raghav
Swapan K. Das · Rakha H. Das · Ved P. Chaturvedi
Hasi R. Das

Association of mannose-binding lectin gene (MBL2) polymorphisms with rheumatoid arthritis in an Indian cohort of case-control samples

Received: 9 June 2005 / Accepted: 11 August 2005 / Published online: 12 October 2005
© The Japan Society of Human Genetics and Springer-Verlag 2005

Abstract Single nucleotide polymorphisms in the mannose-binding lectin (MBL2) gene, as well as the serum MBL2 level, have been associated with various autoimmune diseases. We investigated whether such polymorphisms and/or the serum MBL2 level were associated with rheumatoid arthritis (RA) in an Indian population. The frequency of the B variant (codon 54) of the MBL2 gene was quite frequent in the healthy Indian population and was significantly ($P=6.35\times 10^{-6}$) lower in RA patients. We replicated this association ($P=1.78\times 10^{-5}$) in an independent cohort of control individuals. Promoter polymorphism at -550 nt showed a significant overrepresentation ($P=0.003$) of the minor allele G in severe RA patients compared with the less severe group. Haplotype LYA frequency was significantly ($P=0.03$) high in the less severe group, while the frequency of the HYA haplotype was significantly ($P=0.04$) increased in the severe RA patients. No statistically significant difference in serum MBL2 was observed as a whole, but the individuals homozygous for the LYA haplotype had significantly lower ($P=0.017$) serum MBL2 levels compared with individuals homozygous for the HYA haplotype. Therefore, the B variant of the MBL2 gene may be associated with protection from RA in our study population, and the promoter polymorphism (-550 nt) seems to have some role in disease progression.

Keywords Rheumatoid arthritis · Mannose-binding lectin · Gene polymorphisms · Complement · Indian population

B. Gupta · C. Agrawal · S.K. Raghav · S.K. Das
R.H. Das · H.R. Das (✉)
Institute of Genomics and Integrative Biology,
Delhi University Campus, Mall Road, Delhi 110 007, India
E-mail: hdas@igib.res.in
Tel.: +91-11-27662581
Fax: +91-11-27667471

V.P. Chaturvedi
Department of Rheumatology, Army Hospital,
New Delhi 110010, India

Introduction

Rheumatoid arthritis (RA), an autoimmune disease, is characterized by chronic inflammatory symptoms leading to damage of synovial tissue and joints (Kim-Howard et al. 2005). A higher relative risk of developing the disease in siblings of affected individuals (λ_s) suggests the importance of genetic factors in RA (Suzuki et al. 2003). Among the multiple genes believed to be involved, the HLA-DRB locus is one of the well-studied gene loci reported to be associated with this disease (Suzuki et al. 2003; Weyand and Goronzy 2000). Our group recently found an association of tumor necrosis factor- α microsatellite with susceptibility and progression of RA in an Indian population (Agrawal et al. 2005) that is distinct from other populations (Agrawal et al. 2005; Field et al. 1997). The precise etiology of this disease is unknown, but immune dysfunction is clearly one of its central features. Several autoantibodies are produced in conjunction with manifestations of RA (Goldbach-Mansky et al. 2000). Rheumatoid factors are the characteristic autoantibodies, binding to the Fc region of agalactosylated IgG (IgG₀) molecules (Carson et al. 1987; Corper et al. 1997) and found in elevated levels in the sera of RA patients. IgG₀ has been directly implicated in the pathogenesis of RA (Bond et al. 1996).

Serum mannose-binding lectin (MBL2), an acute phase protein, can induce complement activation by interacting with IgG₀, and the MBL-IgG₀ complex accumulates in the joint fluid of RA patients (Malhotra et al. 1995). MBL2, the active molecule with both collagenous and C-type lectin domain, has a structural similarity (quaternary) with complement protein C1q (Presanis et al. 2003; Turner and Hamvas 2000). Binding of MBL2 to its ligands can activate MBL-associated serine proteases (MASPs), which are homologous to C1q-associated proteases C1r and C1s (Ohta et al. 1990). Therefore, binding of MBL2 to IgG₀ in the sera of RA patients may be one of the reasons for chronic inflammation in RA cases, and recent studies have indicated its

importance in various autoimmune diseases, including RA (Turner and Hamvas 2000; Tsutsumi et al. 2001; Takahashi et al. 2005; Wang et al. 2001). The serum MBL2 level is greatly affected by the polymorphisms of the MBL2 gene (Tsutsumi et al. 2001; Takahashi et al. 2005; Lipscombe et al. 1992; Kilpatrick 2002). The MBL2 gene (locus ID 4153), located on chromosome 10 at position 10q21.1, contains four exons (Sastry et al. 1989). The differential distribution of the normal allele A and the three allelic variants (D, B, and C) due to polymorphism in the codons 52, 54, and 57, respectively, on exon 1 and those in the promoter region at -221 and -550 nt (Fig. 1) has been reported in various populations (Kilpatrick 2002). The variants of the exon 1 cause a change in the amino acid at the motif that is thought to affect the tertiary structure of the collagenous region of the MBL2 protein (Larsen et al. 2004).

Despite several studies, results are conflicting, and no such evidence of an association of MBL2 with RA is available for an Indian population. In our present study, we genotyped these polymorphisms and looked for an association in an Indian cohort of 119 healthy controls and 120 RA patients (including 76 severe and 44 less severe cases). The serum concentration of MBL2 was also determined to check for any association with the polymorphisms and RA. We also replicated our studies after genotyping an independent cohort of 145 healthy control individuals.

Materials and methods

Subjects and samples

Blood samples from 120 patients with RA were collected at the Department of Rheumatology, Army Hospital, Research and Referral, New Delhi, India, after a thorough investigation by the rheumatologists. All patients fulfilled the American College of Rheumatology's classification criteria for the disease. Samples from 119 healthy individuals served as controls (control cohort 1).

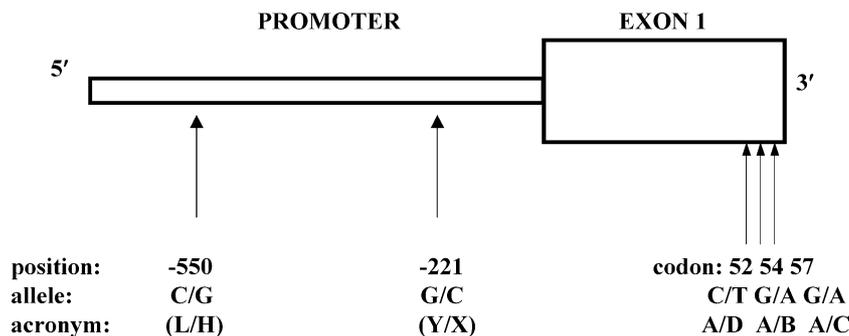
Medical histories and records of the patients, including disease duration, duration of morning stiffness, presence of extraarticular manifestations, presence

of bone deformities, and rheumatoid factor, were collected under the supervision of the rheumatologists. The disease activity score (DAS; Prevoo et al. 1995) was calculated on the basis of the number of tender joints involved, number of swollen joints, erythrocyte sedimentation rate (ESR), and global score. Finally, RA patients were stratified for the severity of their disease based on DAS score and other important medical records. In our cohort, the mean DAS score was 8.15 (range 4.7–9.2), with only one individual with a DAS score < 5.1 (the international cut-off of high disease activity). So for our study, we classified all patients with DAS scores > 8.15 as severe. Patients with DAS scores ≤ 8.15 were classified into severe or less severe groups based on other important medical records. Out of 120 RA patients, 76 were categorized as severe (25 ± 2.51 tender joints, 24 ± 3.8 swollen joints, 137 ± 21.6 ESR, DAS score 8.8 ± 0.35) and 44 as less severe (14.68 ± 5.97 tender joints, 13.86 ± 5.6 swollen joints, 95.6 ± 26 ESR, DAS score 7 ± 0.9). Both the control and case individuals were matched for age, gender, and ethnicity. Age in the control group was 40 ± 5 years and in the RA group was 45 ± 6 years. Considered cases had clinically active disease of consistent duration (8 ± 2 years for severe RA patients and 7 ± 1.5 years for less severe RA patients). Informed consent was obtained from each individual. The institute and hospital ethics committees approved the project.

Samples from control cohort 1 and the patients were collected in two separate aliquots. Aliquots with uncoagulated whole blood samples [ACD-A buffer in sterile vacutainers (Greiner Bio-One)] were used for DNA isolation, and sera separated from other aliquots was used to determine the MBL serum level. The serum samples were stored at -70°C. Genomic DNA was isolated from whole blood following standard procedures (Miller et al. 1988).

For replication of our genotypic association, we ascertained an independent cohort of 145 healthy individuals of 40 ± 5 years of age who were matched for gender and ethnicity (control cohort 2). The DNA samples of these healthy individuals were obtained from our institute's repository. These samples had been collected previously under a population genetics research program with proper informed consent.

Fig. 1 Organization of the promoter region and exon 1 of the human mannose-binding lectin (MBL2) gene. The arrows indicate the positions of single nucleotide polymorphisms (SNPs) resulting in the variant alleles. The graphic is not drawn to scale. Alleles are shown as major/minor alleles



Genotyping

In the patients and in control cohort 1, three exonic polymorphisms (codons 52, 54, and 57) and two promoter polymorphisms (−221 and −550 nt) of the MBL2 gene were genotyped by homogeneous mass EXTEND (hME) assay. Genotypes were confirmed independently by one-direction sequencing. Samples from control cohort 2 were genotyped by direct sequencing only.

Homogenous mass EXTEND (hME) assay

Homogenous mass EXTEND assay based on the annealing of a mass extend primer adjacent to the polymorphic site of interest, extension of the primer through the polymorphic site, and generation of a unique mass product was performed using the vendor's recommended protocol (http://www.sequenom.com/applications/hme_assay.php). For primer sequences, see Table 1. The mass of the extension product was ascertained by SpectroCHIP placed into the MALDI-TOF. Once determined, the genotype was simultaneously called in real time with SpectroTYPER RT software.

Direct sequence analysis

We also reconfirmed our genotype by direct sequence analysis. Promoter variants were sequenced in a 1002-bp PCR amplified product (obtained 600–650 bp clean sequence; Tsutsumi et al. 2001), and exonic variants were sequenced in a 685-bp PCR amplicon (Lipscombe et al. 1992). One-direction sequencing of the purified PCR amplicons was performed with forward primer only, using BigDye Terminator Cycle Sequencing Ready Reaction Kit (version 3.0) following the vendor's recommended procedure. The product of sequencing reaction was further electrophoresed and analyzed using the gene analyzer, ABI PRISM 3730 (Applied Biosystems, MA, USA). The gene sequence obtained was again

analyzed for the polymorphisms using DNASTar and Chromas software.

We finally obtained the genotypes of 120 case individuals, 119 control individuals from control cohort 1, and 90–145 individuals from control cohort 2. Only genotypes with high-quality scores were used for association analysis.

Measurement of the serum MBL2 concentration

Serum concentration of MBL2 was measured by a specific enzyme immunoassay. Mannan-coated microtiter plates were blocked for unoccupied sites using 1% bovine serum albumin (BSA; Sigma Chemicals, MO, USA) for 1 h at 37°C in Tris-buffered saline (TBS, containing 10 mM Tris and 0.9% NaCl, pH 7.2) containing 0.05% Tween-20 (TBST buffer). An amount of 100 µl/well of serum samples diluted 1:16 in TBST containing 50 mM CaCl₂ was incubated in the plates overnight at 4°C. The Ca²⁺ was included because binding of MBL2 is calcium dependent. All samples were analyzed in duplicates. The plates were washed three times with TBST and incubated with 100 µl/well of 1 µg/ml mouse anti-MBL monoclonal antibody (US Biologicals) for 2 h at 37°C. After incubation, the bound antibody was detected using 100 µl/well of 1 µg/ml anti-mouse IgG-HRP conjugate (BD Biosciences). The enzymatic activity of HRP was assessed using TMB (BD Biosciences) as the substrate. The plates were incubated for 20 min at room temperature. The reaction was stopped with the stop solution (BD Biosciences), and optical densities at 450 nm were measured. The standard curve was generated with each assay performed using the serial dilution of the purified MBL (a kind gift of Dr. S. Theil, University of Aarhus, Department of Medical Microbiology and Immunology, Denmark, to Dr. Taruna Madan of our institute). BSA (Sigma Chemicals), a non-sense antigen, was included as a negative control in each assay performed. The wells incubated with TBST instead of serum were used as the blanks.

Table 1 Primers used in genotyping of mannose-binding lectin (MBL2) by homogenous mass EXTEND (hME) assay (*nt* nucleotide)

Position	Primers for hME assay	Amplicon length (bp)
−550 nt, rs11003125	5' ACGTTGGATGGAGAAAATGCTTACCCAGGC 3' 5' ACGTTGGATGCAACCCAGCCAGAATTAAC 3' 5' GCTTACCCAGGCAAGCCTGT 3' ^a	96
−221 nt, rs 7096206	5' ACGTTGGATGACGGTCCCATTTGTTCTCAC 3' 5' ACGTTGGATGTTTCATCTGTGCCTAGACACC 3' 5' CCCATTTGTTCTCACTGCCAC 3' ^a	102
223 nt (codon 52), rs 5030737	5' ACGTTGGATGGTACCTGGTTCCCCCTTTTC 3' 5' ACGTTGGATGGTAGCTCTCCAGGCATCAAC 3' 5' TCCCTTGGTGCCATCAC 3' ^a	99
230 nt (codon 54), rs1800450	5' ACGTTGGATGGTACCTGGTTCCCCCTTTTC 3' 5' ACGTTGGATGGTAGCTCTCCAGGCATCAAC 3' 5' CCTTTTCTCCCTTGGTG 3' ^a	99
239 nt (codon 57), rs1800451	5' ACGTTGGATGAGGCAAAGATGGGCGTGATG 3' 5' ACGTTGGATGATTGCAGAGACAGAACAGCC 3' 5' GGCGTGATGGCACCAAGG 3' ^a	97

^aAssay primer used to determine the genotype in the cases/controls

Statistical analysis

Allelic and genotypic frequencies of the five single nucleotide polymorphisms (SNPs) of the MBL2 gene were calculated. Haplotypes were constructed from genotype data of all five SNPs by using PHASE software (Stephens et al. 2001), which also calculates global *P*-values for haplotype-based case/control association using a permutation test. Statistical differences of allelic and haplotypic frequency between case/control and severe/less severe RA patients were calculated by Fisher's exact test, and genotypic association was assessed by chi-square analysis under additive model. *P*-values equal to or less than 0.05 were considered significant. Due to the exploratory nature of the study, we reported uncorrected *P*-values. Linkage disequilibrium values including *D'* and *r*² were calculated using Haploview software, version 2.05 (Barrett et al. 2005), using the genotype data of the patients and control cohort 1. Markers were also checked for Hardy–Weinberg equilibrium (HWE) using a Hardy–Weinberg program (Guo and Thompson 1992).

Statistical differences in MBL2 levels (μg/ml) between cases and controls and between severe and less severe cases were done by the Mann–Whitney *U*-test (Tsutsumi et al. 2001), and those between genotypes of the five markers were tested either by the Mann–Whitney *U*-test or by the Kruskal–Wallis test (when more than two genotypic groups were compared). Allelic association of those markers with the MBL2 level was performed using the QTL association analysis method as implemented in MENDEL software, version 5.7 (Lange et al. 2001). We also compared statistical differences of MBL2 levels using haplotype-based genotype data (as computed by PHASE) by the Mann–Whitney *U*-test. MEGA2 (<http://www.megasoftware.net>; Kumar et al. 2001) was used to construct neighbor-joining tree-based evolutionary distances between MBL2 haplotypes (Juke-Cantor parameter). The tree was rooted with the MBL2 sequence of the chimpanzee (Bernig et al. 2004).

Results

Allelic and genotypic association

A total of five polymorphic sites of the MBL2 gene were analyzed. The region coding for a part of the collagenous domain of MBL2 is in the exon 1 of the MBL2 gene, in which polymorphisms are reported at the codons 52, 54, and 57 in various populations. In an Indian cohort of 119 healthy controls and 120 RA patients, the allelic and genotypic frequency of polymorphisms at codons 52, 54, and 57 of exon 1 were calculated. We observed a statistically high significant difference in allelic ($P=6.35\times 10^{-6}$, OR = 3.972, 95% CI: 2.068–8.041) as well as genotypic ($P=1.09\times 10^{-6}$, OR = 4.942, 95% CI: 2.442–10.402) frequency of the codon 54 polymorphism between case and control cohort 1. The allele A of codon 54 (at nucleotide 230) was overrepresented in the

control group (frequency = 0.19) compared with cases (frequency = 0.05). This SNP had a significant association even after Bonferroni correction of the *P*-value. There was no significant difference in either genotypic or allelic frequency between case and control cohort 1 for the other four SNPs studied (Table 2). We failed to observe any homozygote in our cohort 1 for minor allele of three exonic polymorphisms studied. The genotype of codon 54 polymorphism that was significantly associated in our case/control cohort was not in HWE ($P=0.007$) in our control cohort 1. Departure from HWE due to genotyping error was checked by

Table 2 Mannose-binding lectin (MBL2) gene polymorphisms and association with rheumatoid arthritis (*HW-P* Hardy–Weinberg *P*-value)

Position		Case	Control 1	Control 2
–550 nt (rs11003125)				
Genotype	CC	51	49	41
	CG	54	54	42
	GG	15	16	17
Allele	C	156	152	124
	G	84 (0.35)	86 (0.36)	76 (0.38)
<i>P</i> -genotypic			0.97	0.67
<i>P</i> -allelic			0.87	0.58
HW- <i>P</i>		1	0.84	0.29
–221 nt (rs 7096206)				
Genotype	GG	60	70	46
	GC	50	43	32
	CC	10	6	12
Allele	G	170	183	124
	C	70 (0.29)	55 (0.23)	56 (0.31)
<i>P</i> -genotypic			0.34	0.4
<i>P</i> -allelic			0.16	0.75
HW- <i>P</i>		1	1	0.14
223 nt (rs 5030737)				
Genotype	CC	100	104	132
	CT	20	15	13
Allele	C	220	223	277
	T	20 (0.08)	15 (0.06)	13 (0.045)
<i>P</i> -genotypic			0.481	0.089
<i>P</i> -allelic			0.499	0.128
HW- <i>P</i>		1	1	1
230 nt (rs 1800450)				
Genotype	GG	106	72	94
	GA	14	47	49
	AA	0	0	2
Allele	G	226	191	237
	A	14 (0.06)	47 (0.2)	53 (0.18)
<i>P</i> -genotypic			1.09×10^{-6}	1.12×10^{-5}
<i>P</i> -allelic			6.35×10^{-6}	1.78×10^{-5}
HW- <i>P</i>		1	0.007	0.164
239 nt (rs 1800451)				
Genotype	GG	110	111	129
	GA	10	8	16
Allele	G	230	230	274
	A	10 (0.04)	8 (0.03)	16 (0.05)
<i>P</i> -genotypic			0.822	0.601
<i>P</i> -allelic			0.825	0.611
HW- <i>P</i>		1	1	1

Parentheses enclose minor allele frequency. *P*-allelic and *P*-genotypic are allelic and genotypic *P*-values, respectively. *Controls 1* and *2* are the control cohorts 1 and 2, respectively

genotyping using two different methods (see “Materials and methods”). No difference was observed between genotypes obtained from either method. Genotypes of all the other polymorphisms were in conformance with HWE.

This genotypic association with codon 54 was replicated (genotypic P value = 1.12×10^{-5} , OR = 4.942, 95% CI: 2.442–10.402) when cases were compared with our independent control cohort 2 (Table 2). In our control cohort 2, we observed only two individuals homozygous for minor allele of codon 54. This control cohort 2 was marginally in HWE, with a Hardy–Weinberg P -value = 0.164, and we observed overrepresentation of heterozygote individuals in this cohort.

The RA patients were stratified on the basis of severity of the disease. Out of the five SNPs studied, the promoter polymorphism at –550 nt (rs11003125) showed a significant (P = 0.003, OR = 2.473, 95% CI: 1.321–4.732) overrepresentation of the minor allele G in severe RA patients (frequency 0.42) compared with the less severe (frequency 0.22) group (Table 3).

Linkage disequilibrium and haplotypic association analysis

We calculated linkage disequilibrium between five SNPs studied. No haplotype block was observed under the confidence interval model (as implemented in Haploview). D' -values between marker 2 (–221 nt) through marker 5 (230 nt) ranged from 1 to 0.86, but the r^2 -values were extremely low (Table 4).

Haplotypes were generated using unphased genotype data of the five SNPs with PHASE software, version 2.0. We observed 10 haplotypes with frequency > 0.01 in our case/control cohort. There was a significant difference (global P = 0.01) in haplotypic frequency between cases and controls (cohort 1). Independent analysis of the distribution of haplotypic frequency of each haplotype was also performed. Haplotype CGCAG (or LYB) was significantly overrepresented (P = 0.002, OR = 3.118, 95% CI: 1.476–7.042) in controls (frequency = 0.11) compared with cases (frequency = 0.035). We also observed marginal significance (P = 0.05, OR = 0.629, 95% CI: 0.393–1.001) in the distribution of CCCGG (or LXA) haplotype in our case/control cohort (Table 5).

Table 3 Mannose-binding lectin (MBL2) gene polymorphisms and association with severity of rheumatoid arthritis (NS not significant)

Position	Frequency ^a		P -value	
	Less severe	Severe	Genotypic	Allelic
–550 nt	0.227	0.421	0.0066	0.003
–221 nt	0.261	0.309	NS	NS
223 nt	0.079	0.085	NS	NS
230 nt	0.079	0.046	NS	NS
239 nt	0.056	0.032	NS	NS

^aMinor allele frequency

Table 4 Linkage disequilibrium analysis using five SNPs in the mannose-binding lectin (MBL2) gene

D'	D'				
	–550	–221	223 (codon 52)	230 (codon 54)	239 (codon 57)
r^2					
–550 nt	1	0.23	0	0.25	1
–221 nt	0.01	1	0.94	0.58	1
223 (codon 52)	0	0.02	1	1	1
230 (codon 54)	0	0.02	0.01	1	0.86
239 (codon 57)	0.02	0.01	0	0	1

Linkage disequilibrium statistics shown as D' - and r^2 -values between the five markers of the MBL2 gene in all the samples tested

Results were similar for our control cohort 2 (data not shown).

In the RA patients, the haplotypes LYA (frequency = 0.32), LXA (frequency = 0.19), and HYA (frequency = 0.22) were more frequently observed compared with seven other haplotypes. The frequency of the LYA haplotype was significantly higher in the less severe patients (P = 0.03, OR = 2.019, 95% CI: 1.068–3.802), whereas the frequency of haplotype HYA was observed to be significantly higher (P = 0.042, OR = 0.503, 95% CI: 0.251–0.978) in the severe RA patients (Table 5).

Phylogenetic relationships between observed haplotypes (except LXB) are shown in Fig. 2.

Genotype-phenotype association analysis

We measured the total MBL2 level in the sera of all case and control individuals. We failed to find any statistically significant difference in total MBL2 level between case (mean MBL2 \pm SD = 1.28 ± 0.81 μ g/ml, median = 1.11 μ g/ml) and control (mean MBL2 \pm SD = 1.23 ± 0.77 μ g/ml, median = 1.06 μ g/ml) or between severe (mean MBL2 \pm SD = 1.17 ± 0.82 μ g/ml, median = 0.9 μ g/ml) and less severe (mean MBL2 \pm SD = 1.27 ± 0.74 μ g/ml, median = 1.09 μ g/ml) RA patients. We also failed to find any allelic or genotypic association with the MBL2 level when we considered each marker separately.

In our haplotype-based association analysis with the MBL2 level (using the best reconstructed haplotype by PHASE), we observed a statistically significant difference (P = 0.017, by Mann–Whitney U -test) in the serum MBL2 level between individuals homozygous for CGCGG or LYA haplotype (mean MBL2 \pm SD = 0.91 ± 0.41 μ g/ml; median = 0.80 μ g/ml, n = 26) and GGCGG or HYA haplotype (mean MBL2 \pm SD = 1.70 ± 1.07 μ g/ml; median = 1.62 μ g/ml, n = 16). Moreover, the CGCGG or LYA haplotype, which had a significantly lower MBL2 level (compared with the HYA homozygotes) was overrepresented in the less severe RA patients (Fig. 3).

Table 5 Haplotype frequencies of five marker haplotypes of mannose-binding lectin (MBL2) in cases/controls and severe/less severe rheumatoid arthritis patients (NS not significant)

Haplotype	Synonym	Frequency		P value	Frequency		P value
		Cases	Controls		Less severe	Severe	
CGCGG	LYA	0.321	0.314	NS	0.414	0.281	0.03
CGCGA	LYC	0.033	0.023	NS	0.038	0.024	NS
CGCAG	LYB	0.036	0.113	0.002	0.046	0.028	NS
CGTGG	LYD	0.048	0.038	NS	0.044	0.039	NS
CCCGG	LXA	0.198	0.126	0.05	0.194	0.190	NS
CCCAG	LXB	0.008	0.021	NS	0.008	0.005	NS
GGCGG	HYA	0.222	0.195	NS	0.148	0.273	0.042
GGCAG	HYB	0.014	0.055	NS	0.021	0.011	NS
GGTGG	HYD	0.026	0.022	NS	0.013	0.026	NS
GCCGG	HXA	0.076	0.072	NS	0.033	0.092	NS

Global *P*-value for case/control haplotype association is 0.01

In silico transcription factor binding analysis

The transcription factor search using the TRANSFAC database (TF SEARCH, available at <http://www.cbrc.jp/research/db/TFSEARCH.html>; Heinemeyer et al. 1998), was carried out to evaluate changes in the transcription factor binding potential due to C/G promoter polymorphism at the position -550 nt of the MBL2 gene. We failed to find any significant difference in transcription factor binding affinity due to this polymorphism.

Discussion

Mannose-binding lectin gene polymorphisms have been studied in many populations for the important role of MBL2 in host defense against infection as well as its function in various autoimmune disorders (Tsutsumi et al. 2001; Lipscombe et al. 1992). The presence of mutations in the collagen region has been shown to correlate with low concentration of human MBL2 (Madsen et al. 1998; Garred et al. 1992). Single nucleotide changes in codons 52, 54, and 57 of exon 1 result in amino acid substitutions

Arg 52 to Cys (C to T at nucleotide 223), forming MBL2 D variant; Gly 54 to Asp (G to A at nucleotide 230), forming B variant; and Gly 57 to Glu (G to A at nucleotide 239), forming C variant. Some phenotypic alteration is also attributable to a polymorphic sequence in the promoter region, particularly at -221 and -550 nt positions (L/H and Y/X alleles, respectively; Madsen et al. 1998, 1995). Stanworth et al. (1998) found no association between MBL2 polymorphisms and RA, whereas Kilpatrick (2002) reported higher MBL2 concentrations in the sera of RA patients compared with healthy controls. Horiuchi et al. (2000) reported that MBL2 gene polymorphisms were not risk factors for systemic lupus erythematosus (SLE) or RA in Japanese individuals, whereas Tsutsumi et al. (2001) found that the frequency of allele B in the patient group was mildly increased in SLE, but with RA, the number of patients homozygous for allele B were slightly increased with no statistical significance. No evidence of an association of MBL2 polymorphisms or serum MBL2 level with RA is available for an Indian population.

In the present study, we attempted to investigate the association of MBL2 gene polymorphisms as well as serum MBL2 level with RA in an Indian case/control cohort.

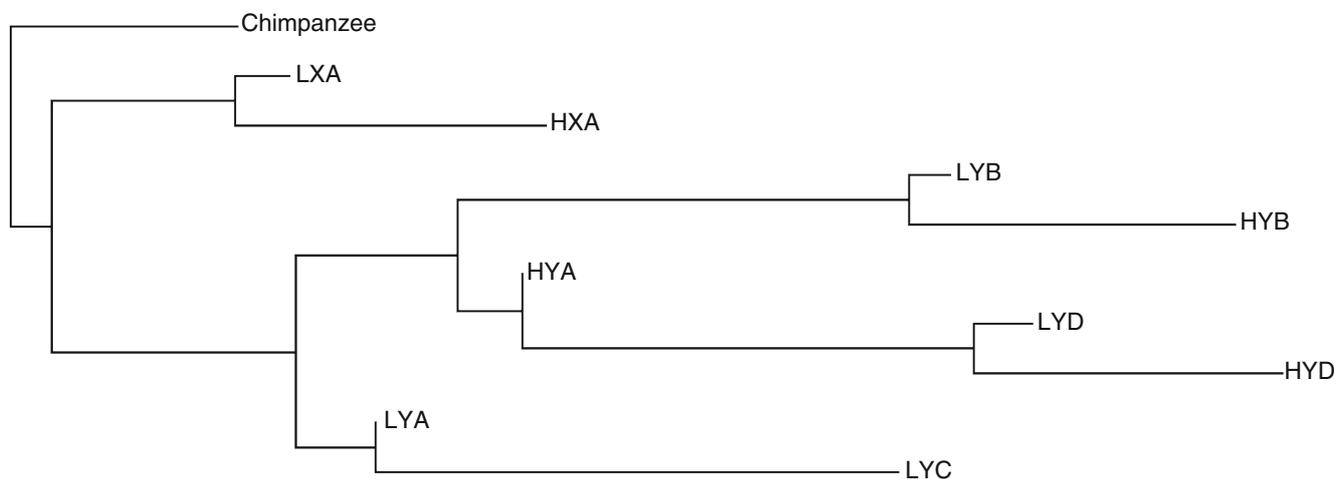
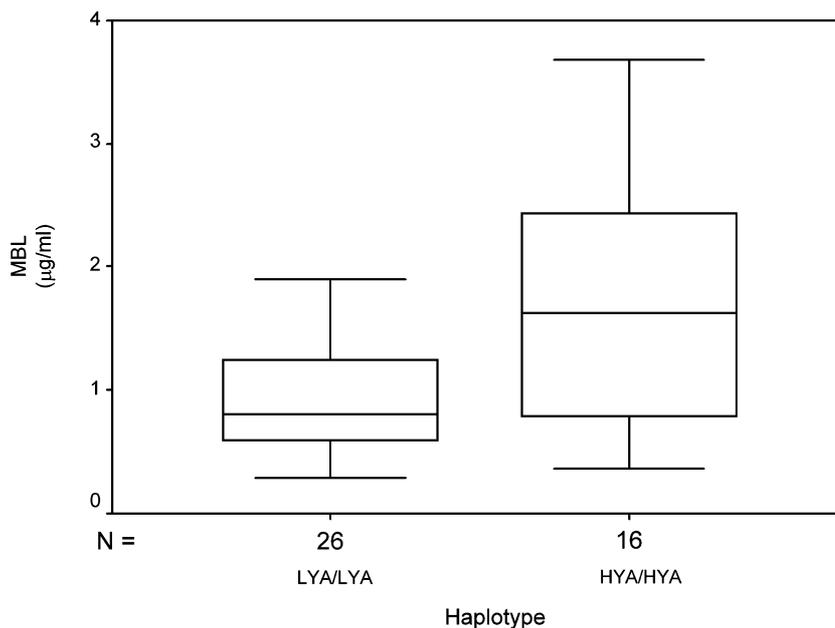


Fig. 2 Neighbor-joining tree (Juke-Cantor correction of nucleotide distances) of the observed haplotypes of mannose-binding lectin

Fig. 3 Relationship between mannose-binding lectin (MBL2) gene haplotypes (LYA/LYA versus HYA/HYA) and serum MBL2 concentration. The *box* represents the interquartile range, which contains the 50% values. The whiskers are *lines* that extend the box to the highest and lowest values. A *line* across the box indicates the median. (*n* number of individuals)



In our study group, mutant allele for codon 54 (A allele, which makes B variant of the MBL2) was significantly overrepresented in the controls (allelic P -value = 3.175×10^{-5} for control cohort 1 and 8.9×10^{-5} for control cohort 2 after Bonferroni correction) compared with patients. In our control cohort 1 ($n = 119$), there was no homozygote for minor allele of codon 54, and we observed only two heterozygotes in our control cohort 2 ($n = 145$). Although the control cohort 2 was in HWE (HW P -value = 0.164), we observed a significant selection of heterozygotes in both control cohorts.

It has been known that replacing Gly with Asp at codon 54 of the MBL2 gene leads to disruption of an oligomerization domain, then leading to formation of functionally inactive lower-order oligomers. These low-order oligomers are unable to activate the lectin complement pathway (Jensen et al. 2005). It has also been known that IgG₀ (agalactosylated IgG), increasingly produced in RA, an autoimmune disease of unknown etiology, can serve as an easy target molecule for MBL2, thereby activating the lectin complement pathway. Synovial fluid has been shown to be rich in the IgG₀-MBL2 complexes, which can then either activate the complement pathway, leading to inflammation, or recruit macrophages to clear the complexes, which secrete proinflammatory cytokines and ultimately lead to inflammatory responses. It may thus be hypothesized that heterozygosity for MBL2 variant allele B, which encodes functionally lower-order oligomers, is associated with protection against RA. This can be well explained by the overrepresentation of the heterozygotes in both of our independent control cohorts. This points toward the selective pressure on the MBL2 gene for this variant (Bernig et al. 2004). Although it is speculative as to what influences have contributed to the preservation of heterozygosity in exon 1, it is likely that changes in the circulating level of higher-order oligo-

mers and the function of variant protein could have a selective advantage in response to environmental pressures. Previous studies have suggested that heterozygotes for B, C, and D could be protected against severe tuberculosis infection (Soborg et al. 2003); therefore, MBL2 deficiency could provide protection against pathogens. Moreover, it is likely that lower levels of functionally active higher-order oligomers, resulting from the structural gene variation at codon 54, could serve to reduce the deleterious effects of excessive complement activation via the lectin complement pathway (Takahashi et al. 2002) hence the hypothesis that in parasite-infested endemic regions, including India, heterozygosity in the MBL2 gene could have a selective advantage. The high incidence of heterozygosity for the structural variants of MBL2, including codon 54, has also been observed in many populations throughout the world, including Africa and South America (Madsen et al. 1998), suggesting the selective advantage of the heterozygotes. Therefore, maintenance of the high heterozygosity in different populations is indicative of a balanced polymorphism system in the MBL2 gene (Bernig et al. 2004).

In our haplotype analysis using genotype information of all five SNPs, haplotype CGCAG (or LYB) was significantly overrepresented in controls. Dropping the genotype information of codon 54 SNP from haplotype analysis decreases the significance of haplotype association (the global P -value drops from 0.01 to 0.49). This leads us to conclude that the haplotypic association comes mainly from the significant association of codon 54 SNP, the minor allele of which is associated with protection against RA.

Rooting the haplotype phylogenetic tree with chimpanzee sequence showed that the LXA haplotype, which is significantly overrepresented in the RA patients, has the evolutionarily oldest background.

In previous population studies in Caucasians and Africans, the MBL2 promoter variants have been described to explain the variations in the MBL2 serum concentrations (Jensen et al. 2005). The HYA, LYA, and LXA haplotypes (haplotypes made by promoter polymorphism -550 , -221 nt, and the three exonic polymorphisms), correlate with different MBL2 concentration in the human serum, with the HYA being the highly secretory haplotype, LYA being the intermediate, and LXA being responsible for the lowest secretion of MBL2 in the serum (Eisen and Minchinton 2003). In our case/control cohort we failed to find any significant difference in distribution of haplotypes constructed with two promoter SNPs (global $P=0.29$). Haplotype HX was absent in Caucasians, Africans, Eskimos, and Japanese, whereas it was present in our cohort in an average frequency of 0.014; data on the association of MBL2 serum level and HX haplotype are lacking.

In our analysis, we observed a significant difference in genotypic ($P=0.007$) as well as allelic frequency ($P=0.003$) of promoter SNP (-550 nt) between severe and less severe RA patients. This association is significant even after Bonferroni correction, although to a lower extent (corrected allelic association $P=0.035$). We also observed a significant difference in five marker-based haplotypic frequency where haplotype LYA was over-represented ($P=0.03$) in the less severe group, while haplotype HYA was overrepresented ($P=0.04$) in severe RA patients. According to Malhotra et al. (1995), the presence of high levels of MBL2-IgG₀ complex in the synovial fluid might trigger a hyperinflammatory response, and, therefore, MBL2 deficiency may have a protective effect against RA. Madsen et al. (1998) found that individuals bearing HY haplotype of MBL2 promoter are the highest secretors and that their serum MBL2 level is also high, whereas individuals with LY haplotype are intermediate secretors. In our genotype-phenotype association analysis, we also observed that individuals homozygous for HYA haplotype have significantly higher MBL2 levels than LYA homozygotes. All of these observations correlate with our observation of the over-representation of HY haplotype in the severe RA patients and of LY in the less severe RA patients of our cohort. Thus, LY haplotype may have some protective effect against degenerative inflammatory changes associated with the severity of RA. Gradual et al. (1998) reported that patients with RA have a larger probability of having undetectable MBL2 serum levels, which leads to a conflicting hypothesis that MBL2 deficiency is associated with susceptibility to RA.

In silico, the transcription factor binding site analysis of these promoter polymorphisms, we also failed to find any significant difference in transcription factor binding when wild allele was replaced by a mutant allele. Thus, the role of the promoter SNPs on the transcription activity of the MBL2 gene and, in turn, its role in disease progression cannot be predicted. More experimental evidence is required to elucidate the role of these promoter polymorphisms in disease severity.

In conclusion, we found that the B variant of codon 54 is quite frequent in the normal healthy Indian population, as is the case in Caucasoids, Chinese, Asians, and Eskimos (Madsen et al. 1995, 1998). This minor allele of codon 54 SNP of the MBL2 gene was associated with protection against RA in our cohort of Indian samples, and the promoter polymorphism at -550 nt (rs11003125) may have some role in disease progression or severity. The presence of other disease severity loci 5' to the MBL2 gene in linkage disequilibrium to this SNP cannot be ignored. The selection pressure on the MBL2 gene and its evolutionary importance can be better explained by a larger population study.

Acknowledgements The authors wish to thank the Council of Scientific and Industrial Research, India, for financial assistance and Mr. A.N. Sirkey of Army Hospital, Department of Rheumatology, for his constant support during the collection of blood samples from the patients. We also thank the Director, Institute of Genomics and Integrative Biology, India, for providing the facility for sequencing.

References

- Agrawal C, Raghav SK, Gupta B, Das RH, Chaturvedi VP, Goswami K, Das HR (2005) Tumour necrosis factor- α microsatellite polymorphism association with rheumatoid arthritis in Indian patients. *Arch Med Res* 36:555-559
- Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263-265
- Bernig T, Taylor JG, Foster CB, Staats B, Yeager M, Chanock SJ (2004) Sequence analysis of the mannose-binding lectin (MBL2) gene reveals a high degree of heterozygosity with evidence of selection. *Genes Immun* 5:461-476
- Bond A, Alavi A, Axford JS, Youinou P, Hay FC (1996) The relationship between exposed galactose and N-acetylglucosamine residues on IgG in rheumatoid arthritis (RA), juvenile chronic arthritis (JCA) and Sjogren's syndrome (SS). *Clin Exp Immunol* 105:99-103
- Carson DA, Chen PP, Fox RI, Kipps TJ, Jirik F, Goldfien RD, Silverman G, Radoux V, Fong S (1987) Rheumatoid factor and immune networks. *Annu Rev Immunol* 5:109-126
- Corper AL, Sohi MK, Bonagura VR, Steinitz M, Jefferis R, Feinstein A, Beale D, Taussig MJ, Sutton BJ (1997) Structure of human IgM rheumatoid factor Fab bound to its autoantigen IgG Fc reveals a novel topology of antibody-antigen interaction. *Nat Struct Biol* 5:374-381
- Eisen DP, Minchinton RM (2003) Impact of mannose-binding lectin on susceptibility to infectious diseases. *Clin Infect Dis* 37:1496-1505
- Field M, Gallagher G, Eskdale J, McGarry F, Richards SD, Munro R, Oh HH, Campbell C (1997) Tumour necrosis factor polymorphisms in rheumatoid arthritis. *Tissue Antigens* 50:303-307
- Garred P, Thiel S, Madsen HO, Ryder LP, Jensenius JC, Svjgarad A (1992) Gene frequency and partial protein characterization of an allelic variant of mannan-binding protein associated with low serum concentration. *Clin Exp Immunol* 9:517-521
- Goldbach-Mansky R, Lee J, McCoy A, Hoxworth J, Yarboro C, Smolen JS, Steiner G, Rosen A, Zhang C, Menard HA, Zhou ZJ, Palosuo T, Van Venrooij WJ, Wilder RL, Klippel JH, Schumacher HR Jr, El-Gabalawy HS (2000) Rheumatoid arthritis associated autoantibodies in patients with synovitis of recent onset. *Arthritis Res* 2:236-243

- Gradual NA, Homann C, Madsen HO, Svejgaard A, Jurik AG, Graudal HK, Garred P (1998) Mannan binding lectin rheumatoid arthritis. A longitudinal study. *J Rheumatol* 25:629–635
- Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* 48:361–372
- Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA (1998) Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acid Res* 26:362–367
- Horiuchi T, Tsukamoto H, Morita C, Sawabe T, Harashima S, Nakashima H, Miyahara H, Hashimura C, Kondo M (2000) Mannose binding lectin (MBL) gene mutation is not a risk factor for systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA) in Japanese. *Genes Immun* 1:464–466
- Jensen PH, Weilguny D, Matthesen F, McGuire KA, Shi L, Hojrup P (2005) Characterization of the oligomer structure of recombinant human mannan-binding lectin. *J Biol Chem* 280:11043–11051
- Kilpatrick DC (2002) Mannan-binding lectin: clinical significance and applications. *Biochim Biophys Acta* 1572:401–413
- Kim-Howard XR, Staudt L, James JA (2005) Update in rheumatoid arthritis therapy. *J Okla State Med Assoc* 98:53–62
- Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17:1244–1245
- Lange K, Cantor R, Horvath S (2001) MENDEL Version 4.0: a complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. *Am J Hum Genet* 69(Suppl):A1886
- Larsen F, Madsen HO, Sim RB, Koch C, Garred P (2004) Disease-associated mutations in human mannose-binding lectin compromise oligomerization and activity of the final protein. *J Biol Chem* 279:21302–21311
- Lipscombe RJ, Sumiya M, Hill AV, Lau YL, Levinsky RJ, Summerfield JA, Turner MW (1992) High frequencies in African and non-African populations of independent mutations in the mannose binding protein gene. *Hum Mol Genet* 1:709–715
- Madsen HO, Garred P, Thiel S, Kurtzhals JA, Lamm LU, Ryder LP, Svejgaard A (1995) Interplay between promoter and structural gene variants control basal serum level of mannan-binding protein. *J Immunol* 155:3013–3020
- Madsen HO, Satz ML, Hogh B, Svejgaard A, Garred P (1998) Different molecular events result in low protein levels of mannan-binding lectin in populations from Southeast Africa and South America. *J Immunol* 161:3169–3175
- Malhotra R, Wormald MR, Rudd PM, Fischer PB, Dwek RA, Sim RB (1995) Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. *Nat Med* 1:237–243
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting-out procedure for extracting DNA from human nucleated cells. *Nucl Acid Res* 16:1215
- Ohta M, Okada M, Yamashina I, Kawasaki T (1990) The mechanism of carbohydrate-mediated complement activation by the serum mannan-binding protein. *J Biol Chem* 265:1980–1984
- Presanis JS, Kojima M, Sim RB (2003) Biochemistry and genetics of mannan-binding lectin (MBL2). *Biochem Soci Trans* 31:748–752
- Prevoo ML, van't Hof MA, Kuper HH, van Leeuwen MA, van de Putte LB, van Riel PL (1995) Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. *Arthritis Rheum* 38:44–48
- Sastry K, Herman GA, Day L, Deignan E, Bruns G, Morton CC, Ezekowitz RA (1989) The human mannose-binding protein gene. Exon structure reveals its evolutionary relationship to a human pulmonary surfactant gene and localization to chromosome 10. *J Exp Med* 170:1175–1189
- Soborg C, Madsen HO, Andersen AB, Lillibæk T, Kok-jensen A, Garred P (2003) Mannose-binding lectin polymorphisms in clinical tuberculosis. *J Infect Dis* 188:777–782
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 68:978–989
- Suzuki A, Yamada R, Chang X, Tokuhiro S, Sawada T, Suzuki M, Nagasaki M, Nakayama-Hamada M, Kawaida R, Ono M, Ohtsuki M, Furukawa H, Yoshino S, Yukioka M, Tohma S, Matsubara T, Wakitani S, Teshima R, Nishioka Y, Sekine A, Iida A, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K (2003) Functional haplotypes of PADI4, encoding citrullinating enzyme peptidylarginine deiminase 4, are associated with rheumatoid arthritis. *Nat Genet* 34:395–402
- Takahashi K, Gordon J, Liu H, Sastry KN, Epstein JE, Motwani M, Laursen I, Thiel S, Jensenius JC, Carroll M, Ezekowitz RA (2002) Lack of mannose-binding lectin-A enhances survival in a mouse model of acute septic peritonitis. *Microbes Infect* 4:773–784
- Takahashi R, Tsutsumi A, Ohtani K, Muraki Y, Goto D, Matsumoto I, Wakamiya N, Sumida T (2005) Association of mannose binding lectin (MBL) gene polymorphism and serum MBL concentration with characteristics and progression of systemic lupus erythematosus. *Ann Rheum Dis* 64:311–314
- Tsutsumi A, Sasaki K, Wakamiya N, Ichikawa K, Atsumi T, Ohtani K, Suzuki Y, Koike T, Sumida T (2001) Mannose-binding lectin gene: polymorphisms in Japanese patients with systemic lupus erythematosus, rheumatoid arthritis and Sjogren's syndrome. *Genes Immun* 2:99–104
- Turner MW, Hamvas RM (2000) Mannose-binding lectin: structure, function, genetics and disease association. *Rev Immunogenet* 2:305–322
- Wang ZY, Morinobu A, Kanagawa S, Kumagai S (2001) Polymorphisms of the mannose binding lectin gene in patients with Sjogren's syndrome. *Ann Rheum Dis* 60:483–486
- Weyand CM, Goronzy JJ (2000) Association of MHC and rheumatoid arthritis. HLA polymorphisms in phenotypic variants of rheumatoid arthritis. *Arthritis Res* 2:212–216