Ulcerative Colitis is Associated with Novel Polymorphisms in the Promoter Region of *MIP-3α/CCL20* Gene

Suck-Chei Choi[1], Eun-Kyung Lee[2], Sungga Lee[2], Soo-Cheon Chae[3], Myeung-Su Lee[1], Geom-Seog Seo[1], Sang-Wook Kim[4], Joo-Jin Yeom[4] and Chang-Duk Jun[2]

[1]Digestive Disease Research Institute, [3]Genome Research Center for Immune Disorders, Wonkwang University School of Medicine, Iksan, [2]Department of Physiology, Kyungpook National University School of Medicine, Daegu, [4]Department of Internal Medicine, Chonbuk National University College of Medicine, Jeonju, Korea

ABSTRACT

Background: We examined global gene expression profiles of peripheral blood mononuclear cells (PBMCs) in patients with ulcerative colitis (UC), and tested whether the identified genes with the altered expression might be associated with susceptibility to UC. **Methods:** PBMCs from 8 UC and 8 normal healthy (NH) volunteers were collected, and total RNAs were subjected to the human 8.0K cDNA chip for the microarray analysis. Real time-PCR (RT-PCR) was performed to verify the results of microarray. One hundred forty UC patients and 300 NH controls were recruited for single nucleotide polymorphism (SNP) analysis. **Results:** Twenty-five immune function-related genes with over 2-fold expression were identified. Of these genes, two chemokines, namely, *CXCL1* and *CCL20*, were selected because of their potential importance in the evocation of host innate and adaptive immunity. Four SNPs were identified in the promoter and coding regions of *CXCL1*, while there was no significant difference between all patients with UC and controls in their polymorphisms, except minor association at g.57A>G (rs2071425, p=0.02). On the other hand, among three novel and one known SNPs identified in the promoter region of *CCL20*, g.−1,706 G>A (p=0.000000055), g.−1,458 G>A (p=0.0048), and g.−962C>A (p=0.0006) were found to be significantly associated with the susceptibility of UC. **Conclusion:** Altered gene expression in mononuclear cells may contribute to IBD pathogenesis. Although the findings need to be confirmed in other populations with larger numbers of patients, the current results demonstrated that polymorphisms in the promoter region of CCL20 are positively associated with the development of UC. **(Immune Network 2005;5(4):205-214)**

Key Words: Inflammatory bowel diseases, ulcerative colitis, crohn's disease, single nucleotide polymorphism

Introduction

The chronic inflammatory bowel diseases (IBDs) ulcerative colitis (UC) and Crohn's disease (CD) remain the most problematic of gastrointestinal disorders. Although the etiologies are unknown, there is evidence linking microbial, immunologic, and genetic factors with pathogenesis of IBD (1-3). Genetic factors seem to have a more dominant role in the pathogenesis of CD than UC, because family rates, concordance rates for monozygotic twins, and concordance rates in offspring for disease type are in general high in CD (4). For example, positive associations with the *CARD15/NOD2* gene on chromosome 16 (IBD1), *OCTN1/SLC22A4-OCT/SLC22A5* genes on chromosome 5q (IBD5), and *DLG5* on chromosome 10q23 have been established in CD (5-7). However, UC also has a high incidence in certain families and shares some susceptibility genes with CD. A linkage between UC and certain regions on chromosomes 3, 7, and 12 has been reported (8). Recently, suggestive linkage to chromosome 2q also has been reported for the UC phenotype (9). However, so far, *MDR1*, a gene of P-glycoprotein, is the only gene identified for 3the susceptibility to the UC pathogenesis (10).

Correspondence to: Chang-Duk Jun, Department of Physiology, Kyungpook National University School of Medicine, 101 Dongin-dong 2-ga, Jung-gu, Daegu 700-422, Korea. (Tel) 82-53-420-4815, (Fax) 82-53-424-3349, (E-mail) cdjun@knu.ac.kr

This research was supported by Kyungpook National University Research Fund, 2005.

Microarray analysis can complement genetic mapping strategies as a means of identifying susceptibility candidate genes in complex diseases. Microarray has been used to study gene expression in mucosal biopsies from patients with UC, patients with CD, and healthy controls (11-13) and identified new categories of genes with altered expression, including *defensin, Reg*, Gro, and calcium-binding *S100* genes. Two recent studies of colonic gene expression in patients with UC and controls also detected increased expression of colon cancer-related *Wnt* signaling and *osteopontin* genes in UC patients (14,15).

Recently, the discovery that *CARD15/NOD2*, a protein primarily expressed in mononuclear cells, could contribute to organ-specific disease placed mononuclear cells and innate immunity at the center of IBD pathogenesis for the first time (5,16,17). In addition, the *Runx 1* gene has been identified as a susceptibility gene for rheumatoid arthritis (18). *Runx 1* is a gene primarily expressed in hematopoietic cells that, when fused to a variety of other genes, plays a role in myeloid malignancies. Polymorphisms in Runx 1-binding sites in the genes encoding SLC22A4, SLC9A3R1, and PDCD1 confer susceptibility to autoimmune diseases, rheumatoid arthritis, psoriasis, and systemic lupus erythematous (18). These data support the notion that hematopoietic cells may play a central role in chronic, inflammatory, and autoimmune conditions.

Here, we therefore evaluated the utility of patients' blood cells, a readily accessible source of material, for the identification of global gene expression pattern in UC because the *NOD2* and *Runx 1* findings implicate mononuclear cells in the pathogenesis of IBD and other autoimmune disorders. We compared UC patients to sex-matched healthy controls to identify genes that were specific to the disease and immune function. Further, to test whether the identified genes with the altered expression in PBMCs could be candidate susceptibility genes to UC pathogenesis, we analyzed single nucleotide polymorphisms (SNPs) of those genes selected. Among 222 genes with over 2-fold expression in UC patients, we identified 25 immune function-related genes. Of these genes, two chemokines, namely, *CXCL1* and *CCL20*, were selected because of their potential importance in the evocation of host innate and adaptive immunity. We scanned the chromosomes of 40 healthy volunteers for the identification of SNPs in the *CXCL1* and *CCL20* genes. Further, we investigated the genotype distributions and allele frequencies of the SNPs in patients with 140 UC patients *vs.* 300 unrelated non-UC healthy controls.

Materials and Methods

Patients and samples. Human peripheral blood mononu-

Table I. Clinical characteristics of the patients with UC at the time of blood drawing

	cDNA microarray[a]		Real time-PCR[b]	
	NH	UC	NH	UC
n	8	8	14	17
Age (yr), mean±SD	32.54±09	41.25±15	26.9±07	38.47±16
[Range]	[26~37]	[24~63]	[21~42]	[20~69]
Sex (m/f)	4/4	4/4	8/6	6/11
Activity Index				
Mild		4		5
Moderate		2		5
Severe		2		7
Disease Location				
Rectosigmoid		6		10
Pancolitis		2		7
Medication				
Non		6		5
5-ASA		2		10
5-ASA+Steroid				2
pANCA				
+				
−		8		17

[a]Sex-matched 8 UC patients and 8 NH controls were recruited for cDNA microarray analysis. [b]Real time-PCR analysis was performed to quantify the expression levels of CCL20 and CXCL1 in patients with UC and NH controls.

clear cells (PBMC) were isolated from normal donors by centrifugation through a discontinuous Ficoll (Amersham) gradient. For microarray analysis, peripheral blood mononuclear cells (PBMCs) from 8 UC patients (4 males and 4 females, aged 41.25±15 years) and 8 normal healthy (NH) controls (4 males and 4 females, aged 32.54±09 years) were obtained from Wonkwang University Medical Center. For real-time PCR (RT-PCR) analysis, the PBMCs were additionally obtained from 17 UC patients (6 males and 11 females, aged 38.47±16 years) and 14 NH (8 males and 6 females, aged 26.9±07 years) controls. Table I presents clinical characteristics of UC patients recruited for microarray and RT-PCR analysis. For SNP studies, blood samples were obtained from 140 UC patients and 300 unrelated non-UC healthy controls. Diagnoses were made by a board-certified gastroenterologist, based on a combination of clinical, radiographic, and endoscopic findings. Clinical characteristics of UC patients were diagnosed according to Truelove and Witts' activity categories (19). The study was approved by the ethics committees of the Medical Faculty of the Wonkwang University, Korea. All participating subjects gave written, informed consent.

RNA isolation. Total RNA was prepared from peripheral blood mononuclear cells (PBMCs) using easy-Blue™ (Intron Biotechnology, Korea) following the manufacturer's instructions. Purity and RNA integrity were evaluated by absorbance at 260 nm and 280 nm. High quality RNAs with A260/A280 ratio over 1.9 and intact ribosomal 28S and 18S RNA bands were utilized for the specific purposes described below. The exact amount of RNA required for a given experiment was range from 15 μg to 20 μg.

cDNA microarray analysis. Total RNA was subjected to reverse transcription while being labeled with either Cy3 or Cy5 by means of 3DNA indirect labeling kit, according to the protocol provided by the manufacturer (Genisphere, Inc., Hatfield, PA). For hybridization, Cy3 and Cy5-labeled cDNAs were applied to a human cDNA chip containing approximately 7500 cDNA clones with 90% known sequences and 10% EST (HSVC V1.0, GenomicTree Inc, Korea). Hybridized slide was scanned with ScanArray 3000. Signal quantification and data processing were performed using Imagene 5.1 and GeneSight 3.5 (BioDiscovery, Inc., El Segundo, CA). Each experiment was repeated three times to reduce a high risk of false positive or false negative results. The results from three independent identical experiments were merged and the merged data were used for subsequent comparisons. The average intensity values of Cy3 and Cy5 for each gene were calculated. Significant differences in gene expression between controls and UC patinets were determined using the Student's *t*-test with a significance threshold of $p < 0.05$. To estimate fold induction, the ratios of Cy5 (treated) to Cy3 (untreated) signal intensities were calculated. The log base 2 of each ratio was determined in order to equalize the magnitude of deflection of up-regulated and down-regulated genes, and gene expression differences were ranked based on absolute values.

Real-time quantitative PCR. CCL20 and *CXCL1* mRNA levels and GAPDH mRNA levels in PBMCs were quantified using a fluorogenic 5'-nuclease PCR assay (20) with a GeneAmp 5,700 sequence detection system (ABI/Perkin-Elmer). For each standard or sample, duplicated reactions containing 2.5 ml of cDNA were incubated for two minutes at 50°C, denatured for 10 min at 95°C, and subjected to 40 cycles of annealing at 55°C for 20 seconds, increasing to 60°C for one minute, followed by denaturation at 95°C for 15 seconds. The gene specific primers used were the same as described above. To detect amplicons generated using the gene specific primers, dual-labeled fluorogenic (Taqman) probes containing FAM (at the 5' end) and TAMRA (at the 3' end) were synthesized (Sigma-Genosys, the Woodlands, Texas, USA). The Taqman probes used included: CCL20, 5'-TGCTAC TCCACCTCTGCGGCGAA-3, *CXCL1*, 5'-TGGTA GCCGC TGGCCGGCGC-3, and *GAPDH*, 5'-ACC CAGAAGACTGTGGATGGCCCC-3'. The number of amplicons in each sample reaction was determined by comparison with a standard curve generated using five 10-fold dilutions of plasmids (10 pg to 1 fg) containing *CCL20*, *CXCL1*, or *GAPDH* cDNA. *CCL20* and *CXCL1* mRNA levels are presented as the number of amplicons per 10^4 GAPDH amplicons.

PCR and DNA sequencing. The coding regions of *CCL20* and *CXCL1* containing the promoter regions were partially amplified by PCR with specific primers using a Perkin-Elmer 9,600 thermal cycler (Perkin-Elmer, Norwalk, CT). PCR was run for 30 cycles of denaturation at 95°C for 30 seconds, annealing at melting temperature of the each primer pair for 40 seconds, and extension at 72°C for 1 minute or 3 minutes. Both sense and antisense strands of PCR products were directly sequenced using the same primers used for the PCR amplification. PCR products purified by PCR purification kit (Millipore, Milford, MA) were used as the template DNA for cycle sequencing. Sequence analysis was performed by Big Dye Terminator cycle sequencing using an ABI 3100 Prism Automated DNA sequencer (Applied Biosystems, Foster City, CA) according to the manufacturer's instruction.

Genotyping. Genotyping was performed by single-base extension (SBE) method with the ABI Prism SnaPshot Multiplex kit (Applied Biosystems). Eight SBE

primers used in this study are listed in Table II. The SBE reaction mixture was prepared according to the manufacturer's instructions. The primer extension reaction was performed at 96°C for 1 minute, followed by 25 cycles at 96°C for 10 seconds, 55°C for 40 seconds, and 60°C for 30 seconds. To clean up the primer extension reaction, 1.5 unit of CIP (New England BioLabs, Boston, MA) was added to the reaction mixture, and the mixture was incubated at 37°C for 90 minutes, followed by 15 minutes at 72°C for purposes of enzyme inactivation. The purified extension products were added to Hi-Di formamide (Applied Biosystems) according to the recommendations of the manufacturer. The mixture was incubated at 95°C for 5 minutes, followed by 5 minutes on ice, and then electrophoresis was performed with the ABI Prism 3,100 Genetic Analyzer. The results were analyzed by the ABI Prism Gene-Scan and Genotyper software (Applied Biosystems).

Statistical analysis. χ^2 tests were used to estimate the Hardy-Weinberg equilibrium. Logistic regression analyses were used to calculate odds ratios (95% con-

Table II. Primer sequences used for genotyping by SBE method

Positions[a]	Primer sequence (5' → 3')
CCL20 gene	
−1706 G>A	GGAACAAGTCTCAGTAAATGTAAAC
−1458 G>A	CCCAGGCTGAACTCCTGACCTCATTACAGGCATGAGCCAC
−962 C>A	CCACTAGACCCAAATAGGTTGAAAGTAAAA
−786 T>C (rs6749704)	CAACAATTCTGAGGCTCTATATTGAGTTATATTAG
CXCL1 gene	
−520 T>C (rs3117604)	TTTGGAGGAAGAGCGTTGTGTCTCGGATTTGTGGCTGGAG
−38 G>T	GGCTCAGCAGGCGGGTCTGGCGGCTGTGAGAGGAG
57 A>G (rs2071425)	TCCCCGGCTCCTGCGAGTGGC
57 A>G (rs4074)	GCCAGGCTGGGGAAACTGCATTCGGA

[a]Calculated from the translation start site.

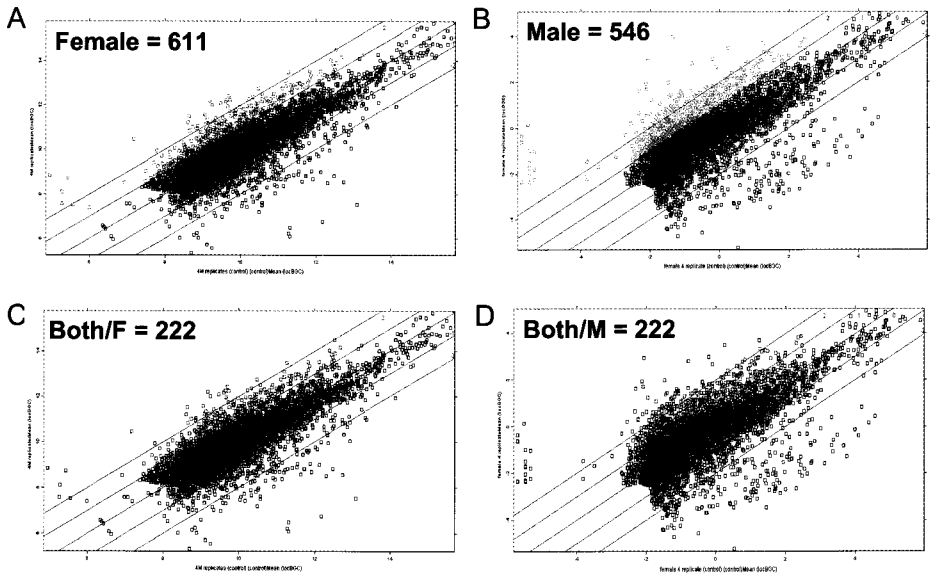

Figure 1. Global gene expression profiles in the PBMC of UC patients and their comparison with NH controls. PBMCs from 8 UC patients (4 m/4f) and 8 NH controls (4 m/4f) were obtained and total RNAs were subjected to the human 8.0K cDNA chip for the microarray analysis. (A) Genes that were over 2-fold upregulated (611) in the female of UC patients compared with female of NH controls. (B) Genes that were over 2-fold upregulated (546) in the male of UC patients compared with male of NH controls. (C) Genes that were over 2-fold upregulated (222) in the both sexes of UC patients compared with female of NH controls. (D) Genes that were over 2-fold upregulated (222) in the both sexes of UC patients compared with male of NH controls. The central diagonals on each scatter plot represent no change from the control state (NH controls). The upper diagonals colored by red (female), green (male), and blue (both sexes) represent each over 2-fold expression changes from the control state.

fidence interval) for SNP sites. Linkage disequilibrium analyses by pairwise comparison of biallelic loci and haplotype and their frequencies were constructed with an EM algorithm or Permutation test with genotyped SNPs. Fisher's exact test or χ^2 test from 2×3 contingency table was applied to analyze the comparison of the frequency of discrete variables between UC patients and unrelated NH controls. A P value of less than 0.05 was considered to indicate statistical significance.

Results

Genes altered in expression in PBMCs from patients with UC as compared to NH controls. The expression levels of each gene from NH controls (labeled with Cy3 before hybridization) *vs.* that of UC (Cy5 labeled) are plotted in Fig. 1. Accordingly, genes whose expression is unchanged lie on the central diagonal while genes whose expression is changed over 2-folds are plotted upper diagonals (Fig. 1). We identified that 611 and 546 genes are over 2-fold expressed in female and male UC patients, respectively. Among these genes, we could extract 222 genes which are over 2-fold expressed in both sexes compared with female or male of NH controls (Fig. 1). Finally, we identified 25 genes whose functions are known to be related with immunity and listed in Table III.

We primarily interested in chemokines because these proteins play central roles in the evocation of host innate and adaptive immunity. Moreover, some of these chemokines, including *CXCL1* and *CCL20*, are known to be expressed in mucosal tissues of UC patients and may possibly involve in the pathogenesis of IBD (21,22). Most interestingly, the chromosomal location of *CCL20* is in 2q33-q37 and this site is well correlated with the recently identified susceptibility loci (2q24-q34 for UC) defined by genome wide meta-analysis (9).

To validate the accuracy of microarray expression profiling data, we performed real-time quantitative PCR Taqman assays. As shown in Fig. 2, the mean values of each *CCL20* and *CXCL1* mRNA levels

Table III. Inflammation-related genes which are up-regulated in the peripheral blood mononuclear cells of patients with Ulcerative colitis[a]

Gene Name	Chromosome locations	Gene ID	Average fold increase[b]
Growth factor receptor-bound protein 7	17q21-q22	H53703	5.328945
Interleukin 18 receptor accessory protein	2q12	AI735624	5.199778
G protein-coupled receptor 30	7p22	AA810225	4.403193
Tumor necrosis factor receptor superfamily, member 14	1p36.3-p36.2	AW073000	3.729934
Neural cell adhesion molecule 2	21q21	AA709271	3.66827
Langerhans cell specific c-type lectin	2p13	AI669693	3.518352
Chemokine (C-C motif) ligand 20	2q33-q37	AI285199	3.363819
A disintegrin and metalloproteinase domain 8	10q26.3	AA279188	3.355435
Fanconi anemia, complementation group G	9p13	AA430675	3.319348
Chemokine (C-C motif) ligand 22	16q13	AI457797	2.922494
Colony stimulating factor 3 (granulocyte)	17q11.2-q12	AI564336	2.904319
Cortical thymocyte receptor (X. laevis CTX) like		AI418753	2.879863
Chemokine (C-X-C motif) ligand 3	4q12-q13	AA935273	2.785816
Activin A receptor, type IB	12q13	AI972229	2.770603
Glucocorticoid receptor DNA binding factor 1	5q31	AA489679	2.729809
Leukocyte receptor cluster (LRC) member 4	19q13.4	AA426216	2.729431
Protein tyrosine phosphatase, non-receptor type 4		AW005820	2.480791
Thyroid hormone receptor-associated protein, 240 kDa subunit	17	AA457462	2.404522
Spi-B transcription factor (Spi-1/PU.1 related)	19q13-q13	N71628	2.396452
Colony stimulating factor 3 receptor (granulocyte)	1p35-p34	AA458507	2.380558
Chemokine (C-X-C motif) ligand 1	4q12-13	W46900	2.362231
Integrin, alpha X (antigen CD11C (p150), alpha polypeptide)	19q13.4	N64384	2.34754
Chemokine (C-C motif) receptor 7	17q12-q21	AI672677	2.315221
Nuclear transcription factor Y, gamma	1p32	AA458991	2.27813
Interleukin 4 receptor	16p12-p11	AA293306	2.202589

[a]Among 222 genes which are 2-fold over-expressed in both sexes of UC patients, as compared to NH controls (Fig. 1), 25 genes whose functions are known to be related with immunity are selected. [b]Data are expressed as fold increase, as compared to the NH controls.

were significantly higher (*CCL20*, 4.6-fold; *CXCL1*, 3.4-fold; $p<0.05$ for both) in the PBMCs of UC patients than in NH controls. This result corroborates that two methods used in the present study reveal overall qualitative agreement.

Polymorphisms in the CCL20 gene of UC patients. The human *CCL20* gene is encoded an open reading frame of 291 bp, consisting of 4 exons and 5-UTR regions separated by intron. To determine the novel molecular variations in the coding and promoter region of human *CCL20* gene, we performed the gene scanning by direct sequencing on genomic DNA samples isolated from 40 unrelated NH controls. We identified a total of 4 SNPs including 3 novel (g.−1,706 G>A, g.−1,458 G>A, and g.−962 C>A) and one known SNPs (g.−786C>T (rs6749704)) in the promoter regions. To precisely determine whether these SNPs are associated to susceptibility of UC in Korean population, we analyzed the genotypes and haplotypes for four SNPs (g.−1706 G>A, g.−1458 G>A, g.−962 C>A, and g.−786 C>T) by SBE method. All genotype frequencies were in Hardy-Weinberg equilibrium (data not shown).

The p values of each SNP were analyzed in UC patients and NH controls (Table IV). The genotype and allele frequencies of g.−1706 G>A in patients with UC were strikingly different from those in NH controls (p=0.000000055 and p=0.000000003, respec-

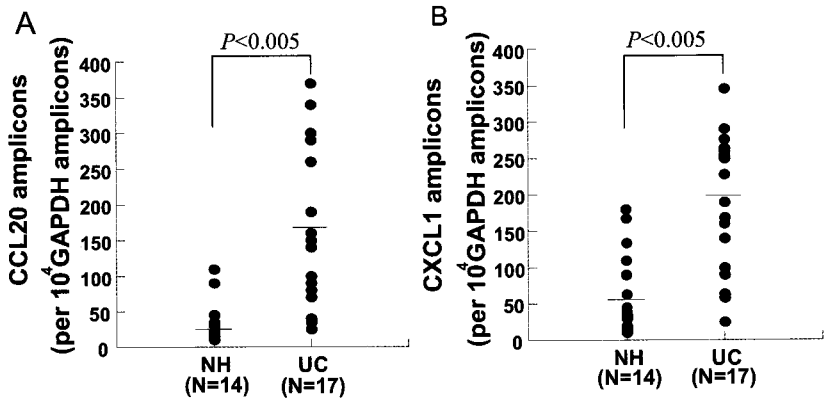

Figure 2. Expression of *CCL20* (A) and *CXCL1* (B) in the PBMCs of UC patients. Real-time quantitative PCR was performed to confirm the results of cDNA microarray as described in *Materials and Methods*. The number of patients included is given below the x-axis. NH: normal healthy controls, UC: Ulcerative colitis patients.

Table IV. Genotype and allele frequencies of the SNP sites in the promoter region of *CCL20* gene

Positions[a]	Genotype/allele	Control n=300 (%)	UC n=140 (%)	Odds ratio[b] (95% CI)	p
g.−1706G>A	GG	114 (38.0)	24 (17.1)	1.00	0.000000055
	GA	137 (45.7)	62 (44.3)	2.15 (1.26~3.66)	
	AA	49 (16.3)	54 (38.6)	5.23 (2.91~9.40)	
	G	365 (60.9)	110 (39.3)	1.00	0.000000003
	A	235 (39.1)	170 (60.7)	2.4 (1.80~3.21)	
g.−1458G>A	GG	10 (03.3)	1 (00.7)	1.00	0.0048
	GA	99 (33.0)	29 (20.7)	2.91 (0.36~23.53)	
	AA	191 (39.7)	110 (78.6)	5.72 (0.73~44.98)	
	G	119 (19.8)	31 (11.1)	1.00	0.0011
	A	481 (80.2)	249 (88.9)	1.99 (1.30~3.04)	
g.−962C>A	CC	174 (58.0)	65 (46.4)	1.00	0.0006
	CA	75 (25.0)	60 (42.9)	2.14 (1.38~3.34)	
	AA	51 (17.0)	15 (10.7)	0.79 (0.41~1.50)	
	C	423 (70.5)	190 (67.9)	1.00	0.432
	A	177 (29.5)	90 (32.1)	1.13 (0.83~1.54)	
g.-786T>C (rs6749704)	TT	173 (57.7)	76 (54.3)	1.00	0.7596
	TC	96 (32.0)	47 (33.6)	1.11 (0.72~1.73)	
	CC	31 (10.3)	17 (12.1)	1.25 (0.65~2.39)	
	T	442 (73.7)	199 (71.1)	1.00	0.417
	C	158 (26.3)	81 (28.9)	1.14 (0.83~1.56)	

[a]Calculated from the translation start site. [b]Logistic regression analyses were used for calculating OR (95% CI; confidence interval).

tively). Most interestingly, compared with NH controls, there was a significant overrepresentation of patients homozygous for the g.−1706G>A allele in UC. A total of 17.1% of UC *vs.* 38.0% of controls were homozygous G, and 38.6% *vs.* 16.3%, respectively, were homozygous A. The overall OR of developing UC for the presence of 2 A alleles was approximately 5-fold increased (OR, 5.23, 95% CI, 2.91-9.40). The genotype and allele frequencies of g.−1458G>A were also significantly different (p=0.0048 and p=0.0011; respectively). The genotype frequencies of g.−962C>A were higher in patients with UC than those with controls (p=0.0006), while allele frequencies of g.−962C>A were not significant (p=0.432; Table IV). These results are strong evidence of which the *CCL20* is associated with the susceptibility to UC.

We next calculated the haplotype frequencies among the g.−1706 G>A, g.−1458 G>A, and g.−962 C>A of *CCL20* gene between UC patients and NH controls. As a result, strong allelic associations (haplotypes) among SNPs were detectable in either the UC patients or NH controls (Table V). Among 8 haplotypes identified in both groups, NH group represented 4 major haplotypes with 93.9% of distribution, while UC group showed 3 major haplotypes with 88.9% of

Table V. The haplotype of CCL20 gene between UC patients and NH controls.

PHFaHaplotypes			HFa		p
g.−1706G>A	g.−1458G>A	g.−962C>A	Control	UC	
G	A	C	37.3	36.1	0.727
A	A	A	23.7	27.4	0.229
G	G	C	17.9	3.2	2.0×10^{-9}
A	A	C	15.0	25.4	0.0002
G	A	A	4.2	4.0×10^{-7}	0.0005
G	G	A	1.4	5.6×10^{-8}	0.046
A	G	C	0.3	3.2	0.0003
A	G	A	0.2	4.7	1.1×10^{-6}

[a]Values were constructed by EM algorithm with genotyped SNPs. HF: haplotype frequency (%).

Table VI. The genotype and allele frequencies of CXCL1 gene between UC patients and NH controls

Positions[a]	Genotype/allele	Control n=300 (%)	UC n=139 (%)	p
−520T>C (rs3117604)	TT	86 (28.8)	29 (20.9)	0.0495
	TC	135 (45.1)	80 (57.5)	
	CC	78 (26.1)	30 (21.6)	
	T	307 (51.3)	138 (49.6)	0.1600
	C	291 (48.7)	140 (50.4)	
−38G>T	GG	127 (42.5)	58 (41.7)	0.5543
	GT	107 (35.8)	56 (40.3)	
	TT	65 (21.7)	25 (18.0)	
	G	361 (60.4)	172 (61.9)	0.1200
	T	237 (39.6)	106 (38.1)	
57A>G (rs2071425)	AA	122 (40.8)	48 (34.5)	0.0200
	AG	112 (37.5)	71 (51.1)	
	GG	65 (21.7)	20 (14.4)	
	A	356 (59.5)	167 (60.1)	0.0100
	G	242 (40.5)	111 (39.9)	
957A>G (rs4074)	AA	63 (21.1)	17 (12.2)	0.0561
	AG	180 (60.2)	98 (70.5)	
	GG	56 (18.7)	24 (17.3)	
	A	306 (51.2)	132 (47.5)	0.8900
	G	292 (48.8)	146 (52.5)	

[a]Calculated from the translation start site.

distribution (Table V). Strikingly, although the estimated haplotype frequencies of two major types (G-A-C and A-A-A) were similar in both groups (UC=63.5%: NH=61.0%), the haplotype frequencies of G-G-C and A-A-C were significantly different between UC and NH (G-G-C: 3.2% and 17.9%, respectively, $p=2\times10^{-9}$; A-A-C: 25.4% and 15.0%, respectively, p=0.0002). Taking all the data together, it could be suggested that the promoter region of *CCL20* gene is highly associated with the susceptibility to UC in Korean population.

Polymorphisms in the CXCL1 gene of UC patients. We identified a total of 4 SNPs, including one novel (g.−38G>T) and three known (g.−520T>C (rs 3117604), g.57A>G (rs2071425), and g.957A>G (rs 4074)) SNPs around the CXCL1 locus mapped on 4q12-q13. We next analyzed the genotype frequencies and allele frequencies between the UC patients and NH controls. The SNPs examined in this study revealed no significant differences, except one minor association at g.57A>G (p=0.02 and p=0.01, respectively), in their genotype and allele frequencies between UC patients and NH controls (Table VI). Accordingly, although there was a limitation in the number of the patients used in this study, i.e., the number of the patients used were relatively small; the typing data suggested that the *CXCL1* gene could not be a major risk factor conferring the susceptibility to UC at least. We further examined allelic associations among 4 SNPs, but found no significant difference between the patients and controls (data not shown). These results suggest that *CXCL1* gene is not associated with the susceptibility to UC in Korean population. Thus, it is conceivable that the *CXCL1* gene could contribute to only inflammatory response in the course of the development of UC, but not participate in the pathogenesis of UC as a genetic factor conferring the predisposition to UC.

Discussion

The exact molecular mechanisms that give rise to UC and CD remain poorly understood. It is believed, however, that in genetically susceptible individuals, both disorders result from an inappropriate or excessive immune response to environmental factors, such as elements of the intestinal microflora (23). Since peripheral blood cells not only form the first line of defense against exogenous putative pathogens, but are also recognized as active participants in inflammatory reactions, increased expression of *CCL20* or *CXCL1* in the PBMCs suggests that these chemokines may positively associated in the pathogenesis of UC. In fact, at least a dozen polymorphic genes are known to regulate IBD by positive association (2,24-29). However, their associations to UC are not significant but predisposed only to the some population (7).

In the current study, we found a striking difference of the genotype and allele frequencies at g.−1706G>A in *CCL20* gene of UC patients compared with NH controls (GG vs. AA, p=0.000000055, OR, 5.23; G *vs.* A, p=0.000000003, OR, 2.4). The genotype and allele frequencies at g.−1458G>A and −962C>A were also significant in patients with UC than those with controls (see Table IV). In addition, the estimated haplotype frequencies resulted in significant difference between UC and NH controls (Table V). So far as we know, the present results demonstrate that *CCL20* is the only chemokine gene which shows a strong association with the susceptibility to UC.

The fact that the allele and *CCL20* −1706AA genotype frequency is clearly overrepresented in UC patients compared in NH controls makes it unlikely that this association occurred by chance. How or whether *CCL20* could be involved in the pathogenesis of UC is presently unknown. However, increased expression of *CCL20* in PBMCs (Fig. 1 and 2) and inflamed colonic tissues (22) of UC patients supports a potential importance of this gene in the pathogenesis of UC. In addition, since *CCL20* is thought to link innate and acquired immunity by attracting immature dendritic cells, effector memory T cells and B cells via CCR6 (30), this gene could be a key regulator of immune function-related diseases, not just for UC but for other immune disorders such as rheumatoid arthritis or asthma. Thus, characterizing the genotype and allele frequencies of these polymorphisms with other autoimmune diseases are one of our on-going projects.

Moreover, a recent genome wide meta-analysis for searching IBD susceptibility loci yielded a strong evidence for linkage to markers on chromosome 2q24-q34 with UC (9). Because region on chromosome 2q is the same region where the *CCL20* gene (2q33-q37) is localized, this finding supports our assumption that the polymorphisms in the promoter region of *CCL20* may confer a genetic predisposition to UC. Further genetic-association studies are needed to clarify in more detail the contribution of other genes which are located in the same chromosomal loci to susceptibility for UC. Although we did not currently investigate the association of *CCL20* with CD because of its lower incidence in Korean population, it will be very interesting whether these polymorphisms in *CCL20* are also predisposed in CD.

In the last 10 years, a global effort involving linkage disequilibrium mapping and positional cloning studies of multiply affected families with inflammatory diseases has uncovered a growing number of susceptibility loci on chromosomes 1, 2, 3, 4, 5, 7,

10, 12, 16, 17, 22, and X (8, 9, 31-37). However, so far, not many specific gene associations with susceptibility to UC have been identified. Microarray can complement genetic mapping strategies as a means of uncovering candidate susceptibility genes in complex diseases. Parallel use of classical genetic strategies with microarray technique may enhance the chance to find genetic susceptibility genes that are predisposed in immune disorders, including UC and CD.

In summary, the identification of *NOD2* as the IBD1 susceptibility gene underscored the significance of mononuclear cells in the pathogenesis of IBD and autoimmunity in general (16). Using microarray analysis, we have detected differential expression of a wide variety of genes in PBMCs from patients with UC as compared to NH controls. Of those differentially expressed genes, we selected two chemokines, namely *CCL20* and *CXCL1*, because of their potential importance in host immunity. We identified that the polymorphisms g.−1706G>A, g.−1458G>A, and −962C>A in the promoter region of the *CCL20* gene are positively associated with UC, thereby suggesting that microarray could be a useful technique to identify genetic susceptibility genes when this technique is combined with classical genetic studies. Further studies are needed to clarify which targets may trigger intestinal inflammation in cases with high expression of *CCL20*.

References

1. Yang H, McElree C, Roth MP, Shanahan F, Targan SR, Rotter JI: Familial empirical risks for inflammatory bowel disease: differences between Jews and non-Jews. Gut 34;517-524, 1993
2. Toyoda H, Wang SJ, Yang HY, Redford A, Magalong D, Tyan D, McElree CK, Pressman SR, Shanahan F, Targan SR: Distinct associations of HLA class II genes with inflammatory bowel disease. Gastroenterology 104;741-748, 1993
3. Tysk C, Lindberg E, Jarnerot G, Floderus-Myrhed B: Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. Gut 29;990-996, 1988
4. Taylor KD, Yang H, Rotter JI: Inflammatory bowel disease. II. Gene mapping. Mol Genet Metab 74;22-44, 2001
5. Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF, Ramos R, Britton H, Moran T, Karaliuskas R, Duerr RH, Achkar JP, Brant SR, Bayless TM, Kirschner BS, Hanauer SB, Nunez G, Cho JH: A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature 411; 603-606, 2001
6. Stoll M, Corneliussen B, Costello CM, Waetzig GH, Mellgard B, Koch WA, Rosenstiel P, Albrecht M, Croucher PJ, Seegert D, Nikolaus S, Hampe J, Lengauer T, Pierrou S, Foelsch UR, Mathew CG, Lagerstrom-Fermer M, Schreiber S: Genetic variation in DLG5 is associated with inflammatory bowel disease. Nat Genet 36;476-480, 2004
7. Cho JH: Advances in the genetics of inflammatory bowel disease. Curr Gastroenterol Rep 6;467-473, 2004
8. Satsangi J, Parkes M, Louis E, Hashimoto L, Kato N, Welsh K, Terwilliger JD, Lathrop GM, Bell JI, Jewell DP: Two stage genome-wide search in inflammatory bowel disease provides evidence for susceptibility loci on chromosomes 3, 7 and 12. Nat Genet 14;199-202, 1996
9. van Heel DA, Fisher SA, Kirby A, Daly MJ, Rioux JD, Lewis CM: Inflammatory bowel disease susceptibility loci defined by genome scan meta-analysis of 1952 affected relative pairs. Hum Mol Genet 13;763-770, 2004
10. Schwab M, Schaeffeler E, Marx C, Fromm MF, Kaskas B, Metzler J, Stange E, Herfarth H, Schoelmerich J, Gregor M, Walker S, Cascorbi I, Roots I, Brinkmann U, Zanger UM, Eichelbaum M: Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. Gastroenterology 124;26-33, 2003
11. Zuo L, Ogle CK, Fischer JE, Nussbaum MS: mRNA differential display of colonic mucosa cells in ulcerative colitis. J Surg Res 69;119-127, 1997
12. Lawrance IC, Fiocchi C, Chakravarti S: Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes. Hum Mol Genet 10;445-456, 2001
13. Warner EE, Dieckgraefe BK: Application of genome-wide gene expression profiling by high-density DNA arrays to the treatment and study of inflammatory bowel disease. Inflamm Bowel Dis 8;140-157, 2002
14. Uthoff SM, Crawford NP, Eichenberger MR, Hamilton CJ, Petras RE, Martin ER, Galandiuk S: Association of ulcerative colitis with the inflammatory bowel disease susceptibility locus IBD2 in non-Jewish Caucasians and evidence of genetic heterogeneity among racial and ethnic populations with Crohn disease. Am J Med Genet 113;242-249, 2002
15. Masuda H, Takahashi Y, Asai S, Takayama T: Distinct gene expression of osteopontin in patients with ulcerative colitis. J Surg Res 111;85-90, 2003
16. Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP, Belaiche J, Almer S, Tysk C, O'Morain CA, Gassull M, Binder V, Finkel Y, Cortot A, Modigliani R, Laurent-Puig P, Gower-Rousseau C, Macry J, Colombel JF, Sahbatou M, Thomas G: Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature 411;599-603, 2001
17. Ogura Y, Lala S, Xin W, Smith E, Dowds TA, Chen FF, Zimmermann E, Tretiakova M, Cho JH, Hart J, Greenson JK, Keshav S, Nunez G: Expression of NOD2 in Paneth cells: a possible link to Crohn's ileitis. Gut 52;1591-1597, 2003
18. Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K: An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. Nat Genet 35;341-348, 2003
19. Truelove SC, Witts LJ: Cortisone in ulcerative colitis; final report on a therapeutic trial. Br Med J 1041-1048, 1955.
20. Holland PM, Abramson RD, Watson R, Gelfand DH: Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. Proc Natl Acad Sci USA 88;7276-7280, 1991.
21. Isaacs KL, Sartor RB, Haskill S: Cytokine messenger RNA profiles in inflammatory bowel disease mucosa detected by polymerase chain reaction amplification. Gastroenterology 103; 1587-1595, 1992
22. Kwon JH, Keates S, Bassani L, Mayer LF, Keates AC: Colonic epithelial cells are a major site of macrophage inflammatory protein 3alpha (MIP-3alpha) production in normal colon and inflammatory bowel disease. Gut 51;818-826, 2002

23. Ahmad T, Satsangi J, McGovern D, Bunce M, Jewell DP: Review article: the genetics of inflammatory bowel disease. Aliment Pharmacol Ther 15;731-748, 2001
24. Klein W, Tromm A, Griga T, Fricke H, Folwaczny C, Hocke M, Eitner K, Marx M, Duerig N, Epplen JT: A polymorphism in the IL11 gene is associated with ulcerative colitis. Genes Immun 3;494-496, 2002
25. Kojima Y, Kinouchi Y, Takahashi S, Negoro K, Hiwatashi N, Shimosegawa T: Inflammatory bowel disease is associated with a novel promoter polymorphism of natural resistance-associated macrophage protein 1 (NRAMP1) gene. Tissue Antigens 58;379-384, 2001
26. Parkes M, Satsangi J, Jewell D: Contribution of the IL-2 and IL-10 genes to inflammatory bowel disease (IBD) susceptibility. Clin Exp Immunol 113;28-32, 1998
27. Frenzel H, Hampe J, Huse K, Mascheretti SM, Croucher PP, Lynch NJ, Siebert R, Platzer M, Schreiber S: Mutation detection and physical mapping of the CD11 gene cluster in association with inflammatory bowel disease. Immunogenetics 53;835-842, 2002
28. Klein W, Tromm A, Griga T, Fricke H, Folwaczny C, Hocke M, Eitner K, Marx M, Duerig N, Epplen JT: Interleukin-4 and interleukin-4 receptor gene polymorphisms in inflammatory bowel diseases. Genes Immun 2;287-289, 2001
29. Low JH, Williams FA, Yang X, Cullen S, Colley J, Ling KL, Armuzzi A, Ahmad T, Neville MJ, Dechairo BM, Walton R, Lench NJ, Jewell DP: Inflammatory bowel disease is linked to 19p13 and associated with ICAM-1. Inflamm Bowel Dis 10;173-181, 2004
30. Hieshima K, Imai T, Opdenakker G, Van Damme J, Kusuda J, Tei H, Sakaki Y, Takatsuki K, Miura R, Yoshie O, Nomiyama H: Molecular cloning of a novel human CC chemokine liver and activation-regulated chemokine (LARC) expressed in liver. Chemotactic activity for lymphocytes and gene localization on chromosome 2. J Biol Chem 272;5846-5853, 1997
31. Hugot JP, Laurent-Puig P, Gower-Rousseau C, Olson JM, Lee JC, Beaugerie L, Naom I, Dupas JL, Van Gossum A, Orholm M, Bonaiti-Pellie C, Weissenbach J, Mathew CG, Lennard-Jones JE, Cortot A, Colombel JF, Thomas G: Mapping of a susceptibility locus for Crohn's disease on chromosome 16. Nature 379;821-823, 1996
32. Cho JH, Nicolae DL, Ramos R, Fields CT, Rabenau K, Corradino S, Brant SR, Espinosa R, LeBeau M, Hanauer SB, Bodzin J, Bonen DK: Linkage and linkage disequilibrium in chromosome band 1p36 in American Chaldeans with inflammatory bowel disease. Hum Mol Genet 9;1425-1432, 2000
33. Hampe J, Lynch NJ, Daniels S, Bridger S, Macpherson AJ, Stokkers P, Forbes A, Lennard-Jones JE, Mathew CG, Curran ME, Schreiber S: Fine mapping of the chromosome 3p susceptibility locus in inflammatory bowel disease. Gut 48; 191-197, 2001
34. Hampe J, Schreiber S, Shaw SH, Lau KF, Bridger S, Macpherson AJ, Cardon LR, Sakul H, Harris TJ, Buckler A, Hall J, Stokkers P, van Deventer SJ, Nurnberg P, Mirza MM, Lee JC, Lennard-Jones JE, Mathew CG, Curran ME: A genomewide analysis provides evidence for novel linkages in inflammatory bowel disease in a large European cohort. Am J Hum Genet 64;808-816, 1999
35. Hampe J, Shaw SH, Saiz R, Leysens N, Lantermann A, Mascheretti S, Lynch NJ, MacPherson AJ, Bridger S, van Deventer S, Stokkers P, Morin P, Mirza MM, Forbes A, Lennard-Jones JE, Mathew CG, Curran ME, Schreiber S: Linkage of inflammatory bowel disease to human chromosome 6p. Am J Hum Genet 65;1647-1655, 1999
36. Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, McLeod RS, Griffiths AM, Green T, Brettin TS, Stone V, Bull SB, Bitton A, Williams CN, Greenberg GR, Cohen Z, Lander ES, Hudson TJ, Siminovitch KA: Genomewide search in Canadian families with inflammatory bowel disease reveals two novel susceptibility loci. Am J Hum Genet 66;1863-1870, 2000
37. Duerr RH: The genetics of inflammatory bowel disease. Gastroenterol Clin North Am 31;63-76, 2002