Identification of Ten Additional Susceptibility Loci for Ulcerative Colitis Through Immunochip Analysis in Koreans

Byong Duk Ye, MD, PhD, Hyunchul Choi, MS,† Myunghee Hong, MS,† Woo Jin Yun, MD, PhD,† Hui-Qi Low, MS,† Talin Haritunians, PhD,‡ Kyung-Jo Kim, MD, PhD,* Sang Hyoung Park, MD,* Inchul Lee, MD, PhD,§ So-Young Bang, MD, PhD,‖ Tae-Hwan Kim, MD, PhD,‖ Hyoung Doo Shin, PhD,¶ Daehee Kang, MD, PhD,** Hee-Shang Youn, MD, PhD,†† Yi Li, PhD,‡‡ Jianjun Liu, PhD,‡‡ Dermot P. B. McGovern, MD, PhD,‡ Suk-Kyun Yang, MD, PhD,* and Kyuyoung Song, PhD†*

Background: Recent genetic association studies identified more than 160 susceptibility loci for inflammatory bowel disease in Caucasian populations, but studies in Asian populations are limited. We have previously reported 3 loci associated with Korean ulcerative colitis (UC).

Methods: Using the Immunochip custom single nucleotide polymorphisms (SNP) array designed for dense genotyping of 186 known disease loci from 12 immune-mediated diseases, we analyzed 705 patients with UC and 1178 controls for 536,821 SNPs (89,057 genotyped and 447,764 imputed) in the discovery stage followed by replication in additional 980 affected individuals and 2694 controls in a Korean population.

Results: We confirmed the associations of 10 known UC risk loci in Koreans: rs76418789 in *IL23R* (combined $P = 1.25 \times 10^{-8}$), rs4728142 in *IRF5* (combined $P = 3.17 \times 10^{-8}$), rs1830610 near *JAK2* (combined $P = 2.28 \times 10^{-9}$), rs1555791 near *TNFRSF14* (combined $P = 1.62 \times 10^{-6}$), rs880790 between *IL10-IL19* (combined $P = 3.73 \times 10^{-6}$), rs10185424 between *IL1R2-IL1R1* (combined $P = 1.54 \times 10^{-4}$), rs6478108 in *TNFSF15* (combined $P = 9.28 \times 10^{-5}$), rs861857 between *UBE2L3-YDJC* (combined $P = 3.05 \times 10^{-5}$), rs1801274 in *FCGR2A* (discovery $P = 1.54 \times 10^{-4}$), and rs17085007 between *GPR12-USP12* (discovery $P = 3.64 \times 10^{-4}$). The percentage of phenotype variance explained by the 13 risk loci (including 3 previously reported loci) was 5.61% in Koreans (on the liability scale, population prevalence = 0.0308%).

Conclusions: Our study increased the number of UC susceptibility loci in Koreans to 13 and highlighted the extensive sharing of genetic risk across populations of UC.

(Inflamm Bowel Dis 2016;22:13–19)

Key Words: ulcerative colitis, genetics, Immunochip, Korean

Ulcerative colitis (UC) and Crohn's disease (CD) are 2 major subtypes of inflammatory bowel disease (IBD). IBD is believed to arise by combined genetic and environmental factors that cause dysregulated mucosal immune responses to the gut flora in genetically susceptible individuals.[1] UC is confined to the colon and rectum with inflammations limited to the mucosa, whereas CD can involve in any part of gastrointestinal tract with transmural inflammation. The incidence of IBD is rapidly increasing in China, Japan, and Korea but is still lower than in the West.[2–4] Of the 163 genetic susceptibility loci to IBD identified

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.ibdjournal.org).

Received for publication July 6, 2015; Accepted July 28, 2015.

From the *Department of Gastroenterology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea; †Department of Biochemistry and Molecular Biology, University of Ulsan College of Medicine, Seoul, Korea; ‡The F. Widjaja Foundation Inflammatory Bowel and Immunobiology Research Institute; Cedars-Sinai Medical Center, Los Angeles, California; §Department of Pathology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea; ‖Department of Rheumatology, Hanyang University Hospital for Rheumatic Diseases, Seoul, Korea; ¶Department of Life Science, Sogang University, Seoul, Korea; **Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea; ††Department of Pediatrics, Gyeongsang National University School of Medicine, Jinju, Korea; and ‡‡Human Genetics, Genome Institute of Singapore, Singapore.

Supported by a Mid-career Researcher Program grant through NRF (National Research Foundation of Korea) to K. Song (2014R1A2A1A09005824) funded by the Ministry of Science, Information & Communication Technology and Future Planning and a Korean Health Technology R&D Project grant to S.-K. Yang (A120176) from the Ministry of Health & Welfare, the Republic of Korea.

The authors have no conflicts of interest to disclose.

Reprints: Kyuyoung Song, PhD, Department of Biochemistry and Molecular Biology, University of Ulsan College of Medicine, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 138–736, Korea (e-mail: kysong@amc.seoul.kr); Suk-Kyun Yang, MD, PhD, Department of Gastroenteology, University of Ulsan College of Medicine, Asan Medical Center, 88, Olympic-ro 43-gil, Songpa-gu, Seoul 138-736, Korea (e-mail: sky@amc.seoul.kr).

DOI 10.1097/MIB.0000000000000584

Published online 22 September 2015.

in Caucasians, two-thirds are shared between CD and UC, and 30 CD-specific and 23 UC-specific loci are identified.[5] However, the disease variance explained by these 163 loci is 13.6% for CD and 7.5% for UC.[5] Recent genome-wide association studies (GWASs) of UC and CD in Korean and Japanese populations suggested that genetic associations for UC tend to overlap among different ethnic groups more than for CD that shows well-established ethnicity dependence.[6–9]

We previously reported confirmation of 3 Caucasian IBD risk loci including the *RNF186-OTUD3-PLA2G2E*, major histocompatibility complex (MHC), and *IRF8* locus through GWAS on UC in Koreans. As the 3 risk loci for UC accounted for only 2.55% of the phenotype variance (population prevalence adjusted) in Koreans, we have used the Immunochip array (Illumina, San Diego, CA) to identify additional UC susceptibility loci in the Korean population. The Immunochip custom single nucleotide polymorphisms (SNP) array is a targeted genotyping array with dense marker coverage across 186 known disease loci identified through GWASs on 12 immune-mediated diseases including UC. This platform would provide an opportunity to identify new UC associations with loci implicated in other autoimmune diseases and for fine-mapping experiments. Here, we report the Immunochip analysis for susceptibility loci of UC in Koreans.

MATERIALS AND METHODS

Study Population

A total of 1685 patients with UC and 3872 unrelated healthy controls, randomly divided into 2 independent cohorts, were included in this study. The clinical characteristics of the study subjects are shown in Table, Supplemental Digital Content 1, http://links.lww.com/IBD/B98. All study participants were of Korean descent. All the patients with UC were diagnosed at the IBD Clinic of Asan Medical Center, Seoul, Korea, on the basis of conventional clinical, radiologic, endoscopic, and histopathologic criteria.[10,11] Patients with indeterminate colitis were excluded from the study. Phenotypic subgroups of UC were determined by 2 independent investigators (S.K.Y. and B.D.Y.) and blinded to the results of the genotype analysis, using the Montreal classification.[12] Briefly, patients with UC were classified into 3 categories including ulcerative proctitis, left-sided UC, and extensive UC on the basis of maximal endoscopic extent. Discovery samples consisted of 461 healthy controls recruited from the University of Ulsan College of Medicine and the Asan Medical Center.[13] In addition, we added another 717 preexisting disease-free Korean controls genotyped by Immunochip that were obtained from T.-W. Kim.[14] Controls for the replication cohort consisted of 1236 healthy controls recruited from the Asan Medical Center, 602 from Seoul National University, 599 from the National Biobank of Korea supported by the Ministry of Health, Welfare and Family Affairs, and 257 from the Gyeongsang National University Hospital, a member of the National Biobank of Korea. This study was approved by the Institutional Review Board of Asan Medical Center, and written informed consent was obtained from all subjects.

Genotyping and Quality Controls

Genotyping of the discovery samples, 707 UC cases, and 469 controls (control cohort 1) was performed using Immunochip at the Cedars-Sinai Medical Center, Los Angeles. We also included additional Immunochip data of 727 Korean control subjects (control cohort 2) obtained from T.-W. Kim.[14] The control cohort 2 samples were genotyped at the University of Queensland Diamantina Institute (Brisbane, Queensland, Australia). In total, discovery cohort consisted of 707 cases and 1196 controls. To control possible batch effects, we used the same genotyping and calling quality metrics; cluster separation (<0.5), call rate <0.99. In total, 21,374 SNPs from UC case cohort, 44,836 SNPs from control cohort 1, and 28,120 SNPs from control cohort 2 were excluded due to poor performance. All SNPs on the X, Y, and mitochondrial chromosomes were excluded. SNPs with a call rate below 98% or with a Hardy–Weinberg equilibrium P value of $< 1.0 \times 10^{-5}$ in controls or with a minor allele frequency < 0.01 were excluded. Similarly, we removed all samples with a genotyping rate less than 96% from further analysis. Six individuals from control cohort 2 were removed due to a genotyping rate less than 96%. We then combined the 707 UC cases and 2 sets of controls. Quality control was performed again in combined samples. In the combined samples, markers with a Hardy–Weinberg equilibrium P value of $<1.0 \times 10^{-5}$ in controls, or with a minor allele frequency <0.01, or with differential missingness between cases and controls ($P < 0.05$) were also excluded from the analysis. Finally, 89,057 SNPs remained for further analysis (see Table, Supplemental Digital Content 2, http://links.lww.com/IBD/B99).

We examined the potential genetic relatedness of 1897 samples based on the identity by descent analysis using PLINK 1.07 software (http://pngu.mgh.harvard.edu/purcell/plink/). For each pair of first or second degree-related samples (PI_HAT $>$ 0.25, IBS >0.8), the sample with the lower genotype call rate was removed. Fourteen samples were removed due to sample duplications and/or close genetic relatedness (2 cases and 12 controls). Subsequently, we performed principal-component analysis (PCA) to detect population outliers and stratification using PLINK 1.07 software and R (v3.2.0; http://www.r-project.org/). All SNPs within 5 distinct regions of long-range linkage disequilibrium (LD), including the human leukocyte antigen region on chromosome 6, inversions on chromosomes 8 and 5, and 2 regions on chromosome 11, were excluded during PCA. First, all the 1883 samples (705 cases and 1178 controls) were analyzed together with 194 reference samples from the International HapMap Project to detect population outliers. Second, PCA was used again to detect population stratification among the cases and controls. The second PCA analysis suggested minimal genetic mismatch between the cases and controls. After the SNP and sample quality control procedures, the final discovery cohort data set included 89,057 SNPs in 705 cases and 1178 controls (average call rate of 99.96%; see Table, Supplemental Digital Content 2, http://links.lww.com/IBD/B99).

Imputation

We used software IMPUTE (v2.0; https://mathgen.stats.ox.ac.uk/impute impute_v2.html) to impute genotype data of untyped SNPs in the Immunochip data.[15] Imputation was performed using the Asian reference panel (JPT + CHB) from the 1000 Genomes Project databases (February 2012 release). Imputed genotypes with a genotype probability of <90%, as well as SNPs with an imputation certainty <80% (based on information scores of imputation results), a minor allele frequency <0.01, and a missing rate of >10% were excluded from further analysis. A total of 447,764 imputed SNPs passed quality control and were combined with 89,057 genotyped SNPs for association analysis.

Replication

After the immunochip analysis, SNPs with $P < 10^{-4}$ were selected for replication. We only chose regions with multiple SNPs having association evidence; the top SNP association signal from each region was then used for validation. We chose genotyped SNPs preferentially over imputed SNPs if there was strong LD ($r^2 > 0.8$) between them, even if imputed SNPs might have lower *P* value. SNPs from the previously reported 3 Korean UC susceptibility loci (the MHC region, *IRF8*, and *OTUD3* loci) were not selected for replication.[7] Of 651 SNPs with $P < 10^{-4}$, 226 SNPs were from the MHC region, 186 SNPs from the *JAK2* locus, and 16 SNPs from the *IRF8* locus. Of the remaining 223 SNPs across 15 distinct loci (1p36.1, 1p36.3, 1p22, 1q32, 2q11, 5q33, 7q21, 7q32, 8p23, 9q31, 10q21, 11q24, 14q24, 22q11, and 22q13), we chose only the top signal from each of 15 loci for validation (see Table, Supplemental Digital Content 3, http://links.lww.com/IBD/B100). The imputation revealed 10 additional top signal SNPs from each of 10 loci that showed $P < 10^{-4}$. Among those, 9 SNPs were from the 9 new loci (1p36.33, 2p22, 2p23, 2p25, 2q31, 12q23, 17p11, 17p13, and 22q12.2) and 1 SNP from an overlapping locus (1q32). All the 10 imputed SNPs were selected for validation as the SNP on 1q32 showed a more significant *P* value than that of a genotyped SNP (see Table, Supplemental Digital Content 3, http://links.lww.com/IBD/B100). Although the top signal in the discovery did not reach $P < 10^{-4}$, we included rs76418789 (*IL23R*) and rs6478108 (*TNFSF15*) in the validation as they were reported to be associated with UC and CD in Caucasians.[5,16,17]

The genotyping of the replication cohort was performed with either iPLEX system (Sequenom, San Diego, CA) at the Analytical Genetics Technology Centre, Princess Margaret Hospital/University Health Network in Toronto, Canada, or the TaqMan genotyping technology in the Applied Biosystems 7900HT Fast Real-Time PCR System according to the manufacturer's suggestion. Validation of rs76418789, rs880790, and rs6478108 was performed using TaqMan SNP genotyping assays.

Statistical Analyses

Single-marker association analyses were performed using the Cochran–Armitage trend test. To calculate the genomic inflation factor, we used a set of 3120 "null" SNPs that are not associated with autoimmune diseases. After quality control, 2124 SNPs were used as null markers to generate a quantile–quantile (Q–Q) plot shown in Figure, Supplemental Digital Content 4, http://links.lww.com/IBD/B101. The Manhattan plot (a scatter plot with negative logarithm P values for the SNP association plotted against the SNP positions) was generated using Haploview (v4.2). Association analyses of the combined samples were performed using Cochran–Mantel–Haenszel test. Two significant thresholds were used to assess the strength of association throughout this study. At loci that have been previously associated with UC in Caucasian populations, the significant threshold applied was $P < 3.76 \times 10^{-4}$ (0.05/133), where 133 is the number of UC loci known to underlie Caucasian IBD. The second significant threshold applied was $P < 5 \times 10^{-8}$ for novel loci. Because our study did not identify any new IBD loci that were not already known to be involved in Caucasian IBD, we applied the significant threshold of $P < 3.76 \times 10^{-4}$ to assess the strength of association. A Breslow–Day test was used to test for heterogeneity between odds ratios (OR) of different sample cohorts. Power of our Immunochip samples for detecting the 133 previously reported UC susceptibility loci in Caucasian population was performed using Quanto software package (Version 1.2.4, http://hydra.usc.edu/gxe/). For each reported SNP, the power for detecting the association at significant threshold of $P < 3.76 \times 10^{-4}$ was calculated based on the reported OR in Caucasian IBD and the allele frequency in Korean population (from this study). For subtype analyses (see Study Population for description of 3 subtypes), we randomly divided the controls into 2 halves and performed the following test: Cases with the subphenotype were compared with half of controls using single SNPs association tests. Cases without the subphenotype were compared with the other half of controls in the same way. Then, the test for heterogeneity of OR was performed.

For the prediction of the evolutionary conservation of a functional variant, the PhastCons46wayPlacental and PhyloP scores were obtained from the UCSC Genome Browser.

Proportion of Variance Explained

The proportion of phenotype variance explained by a number of SNPs is computed based on a liability threshold model and obtained by first calculating Nagelkerke R2 in a multi-SNP logistic regression.[18] Nagelkerke R2 is an analog to R^2 in linear regressions and is subsequently transformed to the proportion of phenotype variance on the liability scale using the following method in reference by Lee et al[19]: Let K be the disease prevalence and P be the proportion of cases in the case/control samples. Suppose f(x) and F(x) are the standard normal distribution density and probability function, respectively; $t = F^{-1}(1\text{-}K)$ and $z = f(t)$. Prop_variance=R2* $\frac{K\times(1-K)}{z^2}\frac{K\times(1-K)}{P\times(1-P)}$.

We assumed a UC prevalence of 0.03087% in the Korean population.[2]

RESULTS

Associations Detected at Previously Identified UC Risk Loci

The workflow of the project is shown in Figure, Supplemental Digital Content 5, http://links.lww.com/IBD/B102. After

stringent quality control measures (see Methods) and PCA (Figures, Supplemental Digital Content 6, http://links.lww.com/IBD/B103), we further analyzed 536,821 SNPs in 705 cases and 1178 controls using trend test in logistic regression. The results of association analysis are shown in a Manhattan plot (Fig. 1). A quantile–quantile plot based on the 2124 null SNPs is shown in Figure, Supplemental Digital Content 4, http://links.lww.com/IBD/B101. The genomic inflation factor (λ_{GC}) was 0.968, indicating that the impact of population stratification was negligible in our study samples.

Of the 27 SNPs selected for replication (see Methods for the details on SNP selection), genotyping of 23 SNPs from 22 loci was successful in an independent set of 980 UC and 2694 control samples. see Table, Supplementary Digital Content 3, http://links.lww.com/IBD/B100 provides complete results from the Immunochip and replication for 27 SNPs. Table 1 highlights 10 SNPs that surpassed our replication threshold for known regions of association to UC, $P < 3.76 \times 10^{-4}$ (see Statistical Analyses in Methods), in the Immunochip or combined analysis of the Immunochip and replication. These SNPs represent the top association signals for each of 10 regions that had 1 or more SNPs meeting criteria for replication of Caucasian loci. Regional plots of Table 1 loci that each contains SNPs with UC association at genome-wide significance $P < 5.0 \times 10^{-8}$ are shown in Figure A, Supplemental Digital Content 7, http://links.lww.com/IBD/B104, (*IL23R*), Figure B, Supplemental Digital Content 7, http://links.lww.com/IBD/B104, (*IRF5*), and Figure C, Supplemental Digital Content 7, http://links.lww.com/IBD/B104 (*JAK2-INSL4*).

The strongest association signal was identified at rs1830610 near the *JAK2-INSL4* locus at 9q24 (OR = 1.35, 95% confidence interval, 1.22–1.49, combined $P = 2.28 \times 10^{-9}$; Table 1, Figure C, Supplemental Digital Content 7, http://links.lww.com/IBD/B104). *JAK2* and *INSL4* genes are located 275.1 kb and 27.5 kb centromeric to rs1830610, respectively. A search for a measurement of interspecies conservation at rs1830610 (1000 bp) in the UCSC Genome Browser database revealed PhyloP and PhastCons score of 0.30 and 0.25, respectively, suggesting that it is evolutionarily conserved. However, we did not find any expression quantitative trait loci associations (data not shown).

The second locus identified was a coding variant of *IL23R* (Gly149Arg), rs76418789, and on 1p31 (OR = 1.75, 95% confidence interval, 1.44–2.13, combined $P = 1.25 \times 10^{-8}$; Table 1, Figure A, Supplemental Digital Content 7, http://links.lww.com/IBD/B104). The third locus identified was at rs4728142 on 7q32 (OR = 1.40, 95% confidence interval, 1.24–1.58, combined $P = 3.17 \times 10^{-8}$) in *IRF5* (Table 1, see Figure B, Supplemental Digital Content 7, http://links.lww.com/IBD/B104). Additional associations were obtained for rs1555791 located ~3.6 kb downstream of *TNFRSF14*, rs880790 located ~14.4 kb upstream of *IL10*, and ~12 kb upstream of *IL19*, rs10185424 located ~18 kb downstream of *IL1R2*, rs6478108 in intron 1 of *TNFSF15*, rs861857 located 36 bp downstream of *YDJC*, rs1801274 in *FCGR2A* (His131Arg),

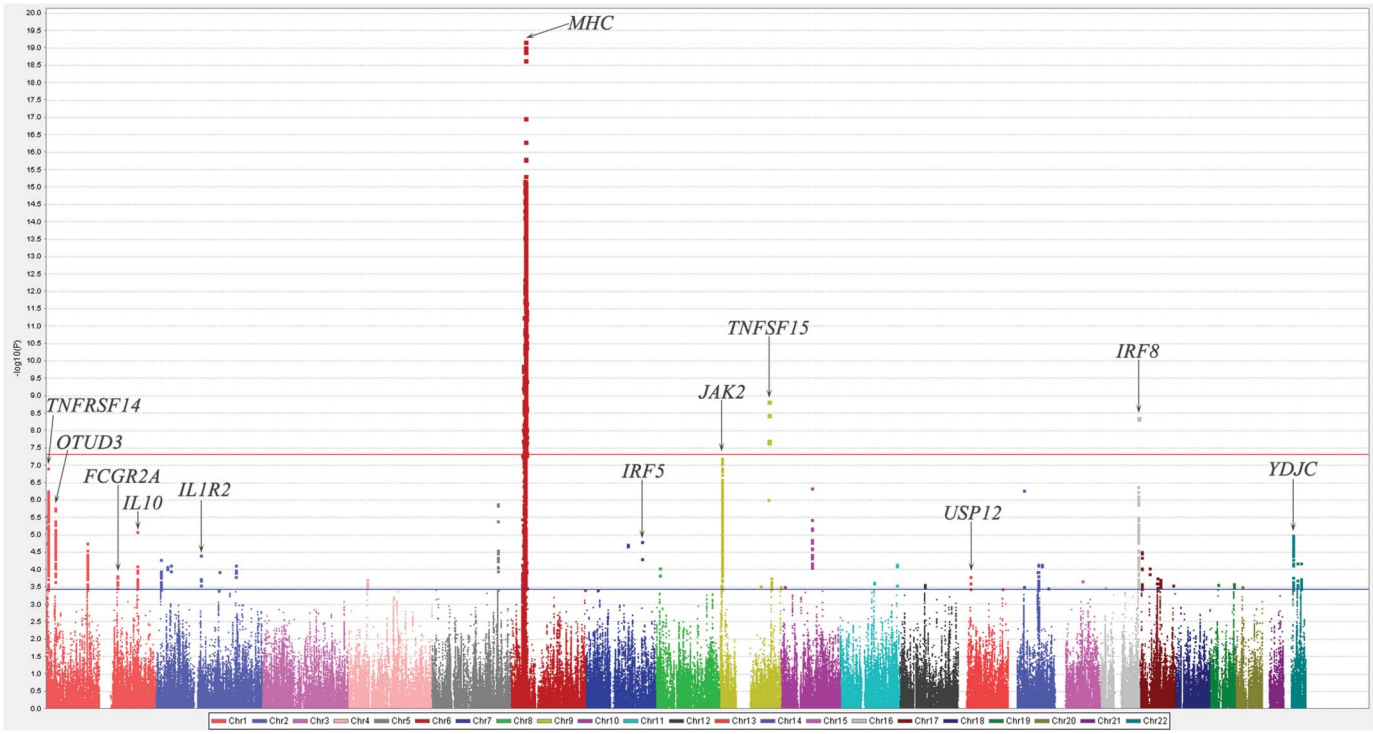

FIGURE 1. Manhattan plot of association results for UC cases and 1178 controls. $-\log_{10}P$ values are shown according to chromosomal location. SNPs within the *JAK2* locus on 9p24, *IRF8* locus on 16q24.1, and the MHC locus showed $P < 5.0 \times 10^{-8}$ (red horizontal line). The blue line indicates significance threshold for replication of SNPs at UC loci established in Caucasians ($P < 3.76 \times 10^{-4}$).

TABLE 1. Association Results of 10 Risk Loci for Ulcerative Colitis in Koreans

SNP	Risk Allele	Chromosome Location	Gene(s)	Study	No. Samples		RAF		P[a]	OR	95% CI	P-BD[b]
					Case	Control	Case	Control				
rs1555791	C	chr:1	*(TNFRSF14)*	Immunochip	705	1174	0.57	0.48	9.01×10^{-7}	1.40	(1.23–1.60)	
		2,498,861		Replication	977	2691	0.55	0.52	2.86×10^{-2}	1.12	(1.01–1.25)	
				Combined[c]	1682	3865	0.56	0.51	1.62×10^{-6}	1.22	(1.12–1.32)	0.0096
rs76418789	G	chr:1	*IL23R*	Immunochip	705	1175	0.95	0.93	3.90×10^{-2}	1.36	(1.02–1.82)	
		67,648,596		Replication	957	2655	0.97	0.93	2.11×10^{-8}	2.11	(1.62–2.76)	
				Combined[c]	1662	3830	0.96	0.93	1.25×10^{-8}	1.75	(1.44–2.13)	0.024
rs1801274	A	chr:1	*FCGR2A*	Immunochip	705	1178	0.81	0.76	1.54×10^{-4}	1.38	(1.17–1.62)	
		161,479,745		Replication	NA							
				Combined[c]								
rs880790	T	chr:1	*(IL10-IL19)*	Immunochip	705	1178	0.75	0.69	4.24×10^{-5}	1.37	(1.18–1.59)	
		206,960,216		Replication	974	2639	0.73	0.70	8.36×10^{-3}	1.17	(1.04–1.31)	
				Combined[c]	1679	3817	0.74	0.70	3.73×10^{-6}	1.24	(1.13–1.36)	0.107
rs10185424	T	chr:2	*(IL1R2)*	Immunochip	705	1177	0.29	0.23	3.78×10^{-5}	1.37	(1.18–1.59)	
		102,662,888		Replication	976	2692	0.27	0.25	1.09×10^{-1}	1.10	(0.98–1.24)	
				Combined[c]	1681	3869	0.28	0.25	1.54×10^{-4}	1.19	(1.09–1.30)	0.026
rs4728142	A	chr:7	*IRF5*	Immunochip	705	1176	0.16	0.11	1.59×10^{-5}	1.55	(1.28–1.88)	
		128,573,967		Replication	978	2692	0.14	0.11	4.61×10^{-4}	1.31	(1.13–1.53)	
				Combined[c]	1683	3868	0.15	0.11	3.17×10^{-8}	1.40	(1.25–1.58)	0.184
rs1830610	T	chr:9	*(JAK2-INSL4)*	Immunochip	705	1178	0.26	0.19	8.97×10^{-8}	1.54	(1.32–1.80)	
		5,260,079		Replication	972	2690	0.23	0.20	8.89×10^{-4}	1.24	(1.09–1.40)	
				Combined[c]	1677	3868	0.24	0.19	2.28×10^{-9}	1.35	(1.22–1.49)	0.033
rs6478108	T	chr:9	*TNFSF15*	Immunochip	705	1177	0.56	0.51	6.62×10^{-3}	1.20	(1.05–1.38)	
		117,558,703		Replication	951	2614	0.55	0.52	3.94×10^{-3}	1.17	(1.05–1.30)	
				Combined[c]	1656	3791	0.56	0.52	9.28×10^{-5}	1.18	(1.09–1.28)	0.730
rs17085007	C	chr:13	*(GPR12-USP12)*	Immunochip	705	1178	0.23	0.18	3.6×10^{-4}	1.33	(1.13–1.57)	
		27,531,267		Replication	NA							
				Combined[c]								
rs861857	G	chr:22	*YDJC*	Immunochip	705	1177	0.49	0.41	1.02×10^{-5}	1.35	(1.18–1.54)	
		21,982,340		Replication	963	2687	0.45	0.43	6.74×10^{-2}	1.10	(0.99–1.23)	
				Combined[c]	1668	3864	0.47	0.42	3.05×10^{-5}	1.19	(1.10–1.29)	0.022

[a]*P* values were calculated using the Cochrane–Armitage trend test.
[b]Asymptotic *P* value of the BD test for heterogeneity of the OR.
[c]Combined *P* value (P_{CMH}) and combined OR were calculated using the Cochran–Mantel–Haenszel test.
(), denotes nearby genes; BD, Breslow–Day test; CI, confidence interval; RAF, risk allele frequency.

rs17085007 located ~109 kb downstream of *USP12*, and ~196 kb upstream of *GPR12*.

Comparison with Caucasian Data

We then compared our data with 133 UC loci recently reported in Caucasians.[5] Our data were available for 117 loci (see Table, Supplemental Digital Content 8, http://links.lww.com/IBD/B105). Significant association of 13 UC loci was found in Koreans including 3 previously reported loci. The lead SNPs at 3 loci (*IRF5*, *FCGR2A*, and *GPR12-USP12*) were exactly the same as those reported in Caucasians. The lead SNPs at 4 loci (*OTUD3*, *TNFRSF14*, *TNFSF15*, and *YDJC*) were in moderate LD ($r^2 > 0.45$) with those reported in Caucasians, and the lead SNPs at 6 loci (*IL23R*, *IL10-IL19*, *IL1R2*, *JAK2-INSL4*, *IRF8,* and *HLA*) were in weak LD ($r^2 < 0.2$). The Caucasian lead SNP rs11209026 of *IL23R* at 1p31 was monomorphic in Koreans, but rs76418789 (Gly149Arg) of *IL23R* showed genome-wide significant association with UC (combined $P = 1.25 \times 10^{-8}$).

We also evaluated the power of our Immunochip samples for detecting the previously reported risk loci. Our samples provided only sufficient power to detect association of the *HLA* locus (power = 0.804) at the significant threshold of $P < 3.76 \times 10^{-4}$ (see Table, Supplemental Digital Content 8, http://links.lww.com/IBD/B105). We also performed SNP–SNP interaction analyses among the 13 loci, including 3 previously reported, and 10 loci identified in this study and found no evidence for interaction in 78 pairs tested (data not shown).

Subphenotype Analysis of the Ten UC Susceptibility Loci

We then examined association between the 10 loci (*IL23R*, *IRF5, JAK2, IL10, TNFRSF14, IL1R2, TNFSF15, YDJC, FCGR2A,* and *USP12*) and UC clinical phenotypes stratified by maximal disease extent. Three loci showed suggestive associations (heterogeneity of odds $P < 0.05$) with extensive colitis: rs1555791 at the *TNFRSF14* locus, rs880790 at the *IL10* locus, and rs6478108 at the *TNFSF15* locus with heterogeneity of odds *P* value of 0.029, 0.016, and 0.006, respectively. However, they did not survive Bonferroni correction (corrected $P = 0.005$; data not shown). None of the SNPs showed significant association with proctitis or left-sided colitis.

DISCUSSION

In our recent GWAS of UC in Koreans, we identified 3 loci including the *RNF186-OTUD3-PLA2G2E* at 1p36.1, the MHC region, and the *IRF8* at 16q24.1, which were all reported in Caucasians previously.[7] In this study, we report 10 additional UC susceptibility loci in Koreans using the Immunochip custom SNP array. In total, 13 UC-risk loci are reported in Koreans, increasing the genetically explained proportion of phenotype variance from 2.55% to 5.61% (see Table, Supplemental Digital Content 9, http://links.lww.com/IBD/B106).

The 10 UC susceptibility loci include 3 loci (*IL23R*, *IRF5,* and *JAK2-INSL4*) with SNPs showing association at $P < 5.0 \times 10^{-8}$ and 7 additional loci at $P < 3.76 \times 10^{-4}$: *TNFRSF14*, *IL10-IL19*, *IL1R2-IL1R1*, *TNFSF15*, *UBE2L3-YDJC*, *FCGR2A*, and *GPR12-USP12*. Identification of additional UC susceptibility loci common to both Caucasians and Asians suggests that common pathways are involved in the UC susceptibility. However, our most strongly associated SNPs at 6 loci were not tightly linked to the previously reported leading SNPs at those loci. It remains to be elucidated how different variants from a given locus might affect pathobiology of UC depending on ethnicity. Compared with the Caucasian lead SNP rs10758669 in the *JAK2-INSL4* locus, our top signal rs1830610 was located closer to *INSL4*. They are 300 kb away from each other, and only weak LD is present between them ($r^2 = 0.2$). In the previous GWAS, we reported promising evidence for association of rs10758669 in *JAK2* with UC (combined $P = 4.49 \times 10^{-5}$). Our Immunochip analysis showed rs1830610 as the lead SNP that was not present in our GWAS chip. No expression quantitative trait loci association was found (data not shown). The significance of rs1830610 in the expression of *JAK2* warrants further study.

Among the 133 reported UC-risk loci, ORs of 94 loci showed consistent direction of effect in Koreans and Europeans. Of those, 32 loci showed suggestive association ($P < 0.05$) in our Immunochip data. Considering the limited power of our study for detecting the previously reported 133 UC risk loci at a *P* value of 3.76×10^{-4}, more loci are expected to be replicated in the Asian population by further studies in larger samples.

In Caucasians, CD and UC were found to share two-thirds of the 163 IBD susceptibility loci. Genes involved in Th17 differentiation including *IL23R, IL12B, JAK2, STAT3, CCR6*, and *TNFSF15* are associated with both CD and UC in Caucasians.[20–22] In our previous study, we did not find significant association of *IL23R* and *TNFSF15* with UC in Koreans.[23] In this study, by expanding our cohort to over 5500 subjects, we found significant or suggestive evidence for association of *IL23R* and *TNFSF15* with UC (Table 1). More overlapping loci between UC and CD could be found, as the sample sizes are increased. Our findings offer new insights into the genetic architecture of IBD and support the complementary value of genetic studies in different populations.

Web Resources

The URLs for the data presented herein are as follows:
The 1000 Genome Project, http://www.1000genomes.org/;
UCSC Genome Browser, http://genome.ucsu.edu/.

ACKNOWLEDGMENTS

The authors would like to thank all participating patients and healthy donors who provided the DNA and clinical information necessary for this study. The authors, Byong Duk Ye, MD, PhD and Hyunchul Choi, MS, contributed equally to this work.

REFERENCES

1. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature*. 2011;474:307–317.

2. Yang SK, Yun S, Kim JH, et al. Epidemiology of inflammatory bowel disease in the Songpa-Kangdong district, Seoul, Korea, 1986–2005: a KASID study. *Inflamm Bowel Dis.* 2008;14:542–549.
3. Ng SC, Bernstein CN, Vatn MH, et al. Geographical variability and environmental risk factors in inflammatory bowel disease. *Gut*. 2013;62:630–649.
4. Zheng JJ, Zhu XS, Huangfu Z, et al. Prevalence and incidence rates of Crohn's disease in main land China: a meta-analysis of 55 years of research. *J Dig Dis.* 2010;11:161–166.
5. Jostins L, Ripke S, Weersma RK, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*. 2012;491:119–124.
6. Asano K, Matsushita T, Umeno J, et al. A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. *Nat Genet.* 2009;41:1325–1329.
7. Yang SK, Hong M, Zhao W, et al. Genome-wide association study of ulcerative colitis in Koreans suggests extensive overlapping of genetic susceptibility with Caucasians. *Inflamm Bowel Dis.* 2013;19:954–966.
8. Yamazaki K, Umeno J, Takahashi A, et al. A genome-wide association study identifies 2 susceptibility loci for Crohn's disease in a Japanese population. *Gastroenterology*. 2013;144:781–788.
9. Yang S-K, Hong M, Zhao W, et al. Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetics susceptibility across ethnic populations. *Gut*. 2014;63:80–87.
10. Lennard-Jones JE. Classification of inflammatory bowel disease. *Scand J Gastroenterol Suppl.* 1989;170:2–6.
11. Park SH, Kim YM, Yang SK, et al. Clinical features and natural history of ulcerative colitis in Korea. *Inflamm Bowel Dis.* 2007;13:278–283.
12. Silverberg MS, Satsangi J, Ahmad T, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol.* 2005;19(suppl A):5–36.
13. Yang SK, Hong M, Choi H, et al. Immunochip analysis identification of six additional susceptibility loci for Crohn's disease in Koreans. *Inflamm Bowel Dis.* 2015;21:1–7.
14. International Genetics of Ankylosing Spondylitis Consortium (IGAS), Cortes A, Hadler J, et al. Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. *Nat Genet.* 2013;45:730–738.
15. Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* 2009;5:e1000529.
16. Yamazaki K, McGovern D, Ragoussis J, et al. Single nucleotide polymorphisms in TNFSF15 confer susceptibility to Crohn's disease. *Hum Mol Genet.* 2005;14:3499–3506.
17. Duerr RH, Taylor KD, Brant SR, et al. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science*. 2006; 314:1461–1463.
18. Wray NR, Yang J, Hayes BJ, et al. Pitfalls of predicting complex traits from SNPs. *Nat Rev Genet.* 2013;14:507–515.
19. Lee SH, Wray NR, Goddard ME, et al. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet.* 2011; 88:294–305.
20. Anderson CA, Massey DC, Barrett JC, et al. Investigation of Crohn's disease risk loci in ulcerative colitis further defines their molecular relationship. *Gastroenterology*. 2009;136:523–529.
21. Franke A, Balschun T, Karlsen TH, et al. Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nat Genet.* 2008;40:713–715.
22. Haritunians T, Taylor K, Targan SR, et al. Genetic predictors of medically refractory ulcerative colitis. *Inflamm Bowel Dis.* 2010;16: 1830–1840.
23. Yang SK, Jung Y, Hong M, et al. No association of TNFSF15 and IL23R with ulcerative colitis in Koreans. *J Hum Genet.* 2011;56:200–204.