creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

Identification of genetic susceptibility loci for intestinal Beh et disease using a genome-wide association study

Yoon Suk Jung

Department of Medicine

The Graduate School, Yonsei University

Identification of genetic susceptibility loci for intestinal Beh et disease using a genome-wide association study

Directed by Professor Jae Hee Cheon

The Doctoral Dissertation

submitted to the Department of Medicine,

the Graduate School of Yonsei University

in partial fulfillment of the requirements for

the degree of Doctor of Philosophy

Yoon Suk Jung

June 2015

This certifies that the Doctoral Dissertation of Yoon Suk Jung is approved.

Thesis Supervisor: Jae Hee Cheon

Thesis Committee Member#1: Dong sik Bang

Thesis Committee Member#2: Young Ho Kim

Thesis Committee Member#3: You Sun Kim

Thesis Committee Member#4: Ji Hyun Lee

The Graduate School
Yonsei University

June 2015

ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to my supervisor, Professor Jae Hee Cheon. He was the person who first introduced me to the field of lower gastrointestinal disease, and trained me to be the gastroenterologist that I am today. His continuous guidance and support enabled me to successfully start and finish this doctorate program. During my doctorate program, he helped guide me into a balanced and insightful-mined researcher.

I also would like to extend my appreciation and thanks to Professor Dongsik Bang, Young Ho Kim, You Sun Kim, and Professor Ji Hyun Lee for sparing their valuable time to advise, guide, and encourage me and also for their comments, and suggestions, which supported the completion of this thesis.

I am also grateful to Seung Won Kim for her kind help and support regarding details of laboratory work. Without his help and support, I could not have completed my dissertation successfully.

Special thanks go to my family for their unconditional love

and sacrifice. Thanks to their constant support and encouragement, I have been able to pursue my dreams.

June 2015

Yoon Suk Jung

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Identification of genetic susceptibility loci for intestinal Beh et disease using a genome-wide association study

Yoon Suk Jung

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Jae Hee Cheon)

Several recent genome-wide association studies (GWAS) have identified susceptibility loci/genes for Behçet's disease (BD). However, no studies have specifically investigated the genetic susceptibility loci associated with intestinal involvement in BD. We evaluated distinctive genetic susceptibility loci/genes associated with intestinal involvement in BD, which can be used to define intestinal BD.

GWAS was performed for 100 Korean BD patients without intestinal involvement, 99 with intestinal involvement, and 557 unaffected individuals. Candidate genes derived from GWAS were then validated using independent cohort samples from 138 patients without intestinal involvement and 196 patients with intestinal involvement. Immunohistochemistry, real-time polymerase chain reaction, enzyme-linked immunosorbent assay, western blot, and gene silencing were performed to evaluate functional gene expression in intestinal epithelial cells and in human and murine colon tissues.

GWAS and validation studies showed intestinal BD-specific

association with the *NAALADL2* gene (rs3914501, OR = 1.914, $P = 3.5 \times 10^{-4}$) and *YIPF7* gene locus (rs6838327, OR = 1.567, $P = 3.4 \times 10^{-4}$). Haplotype analysis was suggestive of associations between *DCAF12, IL10, NAALADL2, PLCB1, SCHIP1,* and *TGFBR3* with intestinal BD and its disease phenotype. The results suggest that these genes are potentially causal variants in association with intestinal BD. Gene functional and pathway analyses indicated that intestinal BD shares inflammatory pathways with inflammatory bowel disease (IBD). *NAALADL2* and *YIPF7* displayed exacerbated inflammatory responses *in vitro* and *in vivo*.

In conclusion, our GWAS results separating between intestinal BD and BD without intestinal involvement enable a more comprehensive analysis of disease specificity and provide insight into the common and different pathogenic mechanisms for intestinal BD and IBD

Key words : Intestinal Behçet's disease, genome-wide association study, pathway analysis, disease susceptible gene

Identification of genetic susceptibility loci for intestinal Beh et disease using a genome-wide association study

Yoon Suk Jung

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Jae Hee Cheon)

I. INTRODUCTION

Behçet's disease (BD) is a rare chronic, inflammatory, and multi-systemic disorder characterized by recurrent oral and genital ulcers, ocular lesions, skin manifestations, arthritis, and vascular, neurological, and intestinal involvement,[1-3] which is prevalent in countries along the old silk route spanning from Far East Asia to the Middle East and the Mediterranean basin. Multiple factors including undefined components of the environment and host genetic changes interact cooperatively to develop the disease,[1-3] Although the nature of these genetic variants remains unknown for the most part, numerous genetic risk factors are thought to contribute to the disease susceptibility,[2-5] In

addition to human leukocyte antigen gene encoding B*51 (*HLA-B*51*) and regions encompassing the major histocompatibility complex (*MHC*) class I, GWAS studies have reported several other BD susceptible non-HLA loci and genes including interleukin (*IL*) 10, IL 23 receptor (*IL23R*)-IL 12 receptor beta 2 (*IL12RB2*), the C-C chemokine receptor 1 gene (*CCR1*), signal transducers and activators of transcription (*STAT4*), genes encoding the killer cell lectin-like receptor family members (*KLRC4-KLRK1*), and the endoplasmic reticulum aminopeptidase gene (*ERAP1*). [2-6]

Intestinal BD (BD with intestinal involvement) is diagnosed if there is a typically shaped ulcer in the gastrointestinal tract and clinical findings meet the diagnostic criteria for BD.[7] Interestingly, intestinal BD is rare in Mediterranean BD patients (0 to 3%),[8, 9] whereas it is relatively frequent in East Asia, including Korea and Japan (5-25%).[10-12] Given these geographic and racial differences in incidence of BD with or without intestinal involvement, intestinal BD is likely to have distinguishable pathogenic pathways from BD without intestinal involvement and there may be distinctive genetic susceptibility loci specifically associated with intestinal BD. Moreover, some phenotypic overlaps (inflammation in the eyes, skin, and intestine), shared genetic associations in the MHC class I region, *IL10,* and *IL23R* and the effectiveness of tumor necrosis factor (TNF)-α blockade therapy implicate common pathogenic pathways between BD and inflammatory bowel disease (IBD).[13, 14] In particular, given that IBD and

intestinal BD share a number of clinical phenotypes in terms of intestinal manifestations, therapeutic applications, and clinical courses,[14] there may be a shared inherited susceptibility between intestinal BD and IBD. Otherwise, intestinal BD and IBD are considered two distinct diseases and thus there may be specific genetic markers for intestinal BD, but not for IBD. Although through several recent GWAS studies investigating BD patients, much is now known about genetic susceptibility loci associated with BD development,[2-6] all of these studies have focused on BD itself and exploring genetic factors associated with intestinal involvement of BD has not been challenged to date because of its rarity.

In the current study, we assessed specific genetic susceptibility loci associated with intestinal BD using a GWAS in the Korean population, which might define specific genetic markers for intestinal involvement of BD. We also investigated the functional consequences of the genetic factors influencing both the disease susceptibility and phenotypes of BD with or without intestinal involvement.

II. MATERIALS AND METHODS

1. Study subjects and DNA Extraction

A total of 533 BD patients of Korean descent, including 238 BD cases

without intestinal involvement and 295 with intestinal BD cases were enrolled in the Behçet's disease clinics of Yonsei University College of Medicine, Severance Hospital, Seoul, Korea between June 2006 and August 2013. Intestinal BD was diagnosed according to the established criteria based on colonoscopic features and clinical manifestations.[7] Only patients who were finally classified as "definite" or "probable" types were included in this study. The demographic and clinical characteristics of the study subjects are summarized in Supplementary Table 1. We obtained the genotyping results in randomly selected 600 individuals matched up sex with the cases from 10,000 healthy Koreans belonging to the Korean Association Resource Project using the same platform. This study was approved by the Institutional Review Boards of Severance Hospital, Yonsei University and all patients and controls provided a written consent. Genomic DNA was extracted from the whole blood samples using the DNA blood maxi kit from Qiagen (Santa Clarita, CA, USA).

2. Genotyping and quality control in GWAS.

A total of 199 patients with BD (100 BD patients without intestinal involvement and 99 with intestinal BD) entered into genotyping using Affymetrix Whole-Genome-Wide Human SNP Array 6.0. (Affymetrix, Santa Clara, CA, USA), in which all samples showed call rates of > 95% finally entered into case–control analysis. After excluding controls with < 95% call

rates, mismatched sex, and subjects who were the potential relatives, a total of 557 control samples of 600 Korean individuals entered into case–control analysis.

In quality controls, SNPs showing *P* value > 1×10^{-4}, if they were not mapped to autosomal chromosomes in cases or controls, those showing < 1% minor allele frequency, or those showing *P* value < 1×10^{-7} Hardy-Weinberg equilibrium (HWE) in controls were excluded from the analysis. The quantile–quantile plot (Q–Q plot) was obtained to evaluate the overall significance of the genome-wide associations and the potential impact of population stratification. After quality control and filtering processes, genotype data for 651,605 SNPs in all BD patients (BD with and without intestinal involvement) vs. healthy controls, 646,355 SNPs in BD patients without intestinal involvement vs. healthy controls, 652,294 SNPs in intestinal BD patients vs. healthy controls excluding SNPs significant in BD patients without intestinal involvement, 637,816 SNPs in patients with intestinal BD patients vs. BD patients without intestinal involvement were used for the final GWAS analysis.

Principal component analysis (PCA) was used to assess the population structure stratification using the software package EIGENSTRAT v.5.0.1[15] with default parameters. For the initial PCA, our Korean samples (199 cases and 597 controls) were analyzed along with unrelated European (CEU) and African (YRI) individuals obtained from 1000 Genomes Project

(phase1_release_v3.20101123) using 391,001 markers which showed > 95% call rates, > 5% MAF, and > 0.05 HWE *P* value. In addition, 199 BD cases were compared with 597 controls. We generated a Manhattan plot of $-\log_{10} P$ using Haploview (V.4.2) (http://www.broadinstitute.org/haploview).

3. Imputation

Aiming to increase our coverage of common variants and capture additional association signals in our study, missing genotypes were imputed in 1000 Genomes Project data released March 2012 (Phase I, version 3) as the reference using IMPUTE (v.2.0) (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html). Reference haplotypes included data from 246 African, 181 American, 286 Asian, and 379 European individuals because combination panel may improve imputation performance.[16] A total of (5,418,765) SNPs with an estimated imputation accuracy of > 0.9 and minor allele frequency (MAF) of > 0.01 were included in the association analysis.

4. SNP Selection and Validation Study

Following the genome-wide discovery analysis, 47 SNPs (25 loci) from regions containing more than 2 SNPs of *P* value $< 1 \times 10^{-4}$ within 100 Kb or lowest *P* value in our GWAS and 13 SNPs (8 loci) previously reported SNPs in association with BD were selected for the replication study in a non-

overlapping set of 295 intestinal BD and 238 BD without intestinal involvement, and 391 controls for whom DNAs were available.

The validation of rs17856199, rs1800871, and rs4500591 was performed using TaqMan genotyping assays (Applied Biosystems, CA, USA) and the validation of the others was performed using BioMark TM HD system (Fluidigm, CA, USA) according to the manufacturer's protocol and recommended quality control measures. The concordance rate in 44 SNPs among duplicated samples between Affymetrix Whole-Genome-Wide Human SNP Array 6.0 and TaqMan genotyping assays or BioMark TM HD system was > 99.9%, indicating a negligible possibility of genotype error.

5. Haplotypic Analysis

To clarify the susceptibility region and narrow down the candidate region, haplotypic associations were analyzed using a method developed by Lake et al.[17] The method is implemented in R (http://www.r-project.org) and provides estimates and the significance of the relative effects of haplotypes on the trait compared to the effects of the baseline haplotype. In this study, a haplotype consisting of a combination of the major alleles from each locus was used as a baseline.

6. In silico analyses of SNP functions

For the Cis- expression Quantitative Trait Locus (cis-eQTL) analysis for

revealing the genetic loci in the control of specific changes in gene expression, we looked the publicly available sources from eQTL Browser (http://www.sanger.ac.uk/ resources/software/genevar/). The cis associations of each SNP and its proxies with the expression of nearby genes were searched in fibroblasts, lymphoblastoid cells, T-cells, and tissues from adipose and skin.[18, 19]

Regulatory SNPs were surveyed using RegulomeDB[20] (http://regulomedb.org) and HaploReg v2 resource (http://compbio.mit.edu/HaploReg). Variants in RegulomeDB analysis can be classified into one of six RegulomeDB categories with scores ranging from 1 to 6 indicating putative functions and a total of 45 SNPs were used for the functional annotation analysis of established prostate cancer risk alleles. The HaploReg analysis was used to identify biological features in sequences containing the BD risk variants and SNPs in high LD (r^2 0.8 in Europeans from the 1000 Genomes Project) and evolutionarily conserved regions based on SIte-specific PHYlogenetic (SiPhy) analysis.[21, 22] Functional elements located in the same regions as index and the correlated SNPs were identified in ENCODE[23] and annotated with potential effects on regulatory motifs based on existing databases such as TRANSFAC, JASPAR, and PBMC.[22] In HapoReg analysis, we selected ASN (CHB+JPT) as the population for LD calculation, and we chose all SNPs in the ASN population as the background for enhancer enrichment analysis.

Prediction of the possible impact of an amino acid substitution on protein function was done by using SIFT (Sorting Intolerant From Tolerant, http://siftdna.org/www/SIFT_dbSNP.html) or polymorphism phenotyping v2 (PolyPhen-2) algorithms, http://genetics.bwh.harvard.edu/pph2), in which "Damaging" or "Probably/Possible damaging" were considered to predict an effect on protein function, respectively.

7. Immunohistochemical staining

For immunohistochemical analysis, formalin-fixed paraffin-embedded (FFPE) tissue sections (21 normal tissues obtained from cancer patients and 28 inflamed tissues from intestinal BD patients after surgical resection were deparaffinized, washed with distilled water, and heated in 10 mM citrate buffer (pH 6.0) for 10 min for antigen retrieval. Sections were then washed with distilled water and treated with 3% H_2O_2 for 5 min. After washing with Tris-buffered saline containing 0.1% Tween-20 (TBS-T), sections were blocked with 2.5% normal horse serum in TBS-T for 1 h and then incubated with NAALADL2 antibody (1:100, Santa Cruz Biotechnology, Santa Cruz, CA, USA) and YIPF7 (1:100, Novus Biologicals, Littleton, CO, USA) diluted in 1% normal horse serum at 4°C overnight. After washing, sections were incubated with biotin-conjugated IgG (1:500, Vector Laboratories, Burlingame, CA, USA) and then treated with reagents from a Vecta-Elite streptavidin-peroxidase kit (Vector Laboratories) with a benzidine substrate

for color development. Sections were counterstained with diluted hematoxylin and examined by light microscopy (Olympus BX41; Olympus Optical, Tokyo, Japan). To quantify gene expression, we studied randomly selected fields for each sample at 100× magnification and scored ranging from 0 to 3.

8. Cell culture, treatment, and enzyme-linked immunosorbent assay (ELISA)

The HT-29 cell line (Korean Cell Line Bank, Seoul, Korea) was maintained at 37℃ in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% antibiotics in a humidified atmosphere of 5% CO_2. Cells were incubated with LPS (1 g/mL) for the indicated time.

Knockdown of a specific gene was achieved by 24 h transfection of siRNA or non-targeting control (AccuTarget, Bioneer, Daejeon, Korea) into HT-29 cells. To assess inflammatory response, cell culture medium was replaced with medium containing LPS (1–2 g/mL) at 24 h post transfection and harvested at 3 h for qRT-PCR analysis, 24 h for immunostaining, and at 72 h for ELISA after LPS-treatment. ELISA was performed using BioLegend human TNF-α ELISA MAXTM kit (Biolegend, San Diego, CA, USA).

9. Quantitative Real-time Reverse-transcription Polymerase Chain Reaction (RT-PCR)

Total RNA was extracted using TRIzol Reagent (Life Technologies, Grand Island, NY, USA) and RNA was reverse-transcribed using SuperScript First-Strand Synthesis kit (Life Technologies) according to the manufacturer-recommended protocol. The cDNAs were mixed in triplicate using SYBR Green master mix (Life Technologies) and pairs of primers (4 pmol of each primer). PCR was done using primers for *NAALADL2, YIPF7*, *TNF-α*, *IL12B*, and *IL22* (AccuTarget qRT-PCR primer, Bioneer). Samples were amplified in a StepOne Plus real-time PCR system (Applied Biosystems) for 40–45 cycles using the following PCR variables: 95℃ for 30 sec, 60–62℃ for 30 sec, and 72℃ for 40 sec. Finally, quantitative analysis was performed using the relative standard curve method and the results were reported as the relative expression or fold change compared to the calibrator after normalization of the transcript level to the endogenous control, *GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE* (*GAPDH*) *or β- ACTIN* (*β-actin*).

10. Trinitrobenzenesulfonic acid (TNBS)-induced colitis of mice

We used 8-week-old male C57BL/6 (Orient, Seongnam, Korea). The mice were maintained on a 12:12-hour light:dark cycle under specific pathogen-free conditions. Mice were slightly anesthetized with an i.p. injection of zoletil (20 mg/kg) and xylazine (10 mg/kg). Colitis was induced by the administration of 5% (w/v) TNBS solution (100 μl) in 50% ethanol into the colon of lightly anesthetized mice via a thin round-tip needle equipped with a

1-mL syringe inserted into the colon 4 cm proximal to the anus. The animals were kept in a head-down vertical position for 2 min to distribute the agents within the entire colon and caecum (day 0). Throughout the experiment, mice were monitored for body weight loss and overall mortality. At day 3, the mice were sacrificed and the colon was quickly removed from the proximal rectum, close to its passage under the pelvisternum, opened longitudinally and gently cleared of stool by PBS. The colon length was measured between the ileo–cecal junction and the proximal rectum. The colon tissue was then used for histological and qRT-PCR analysis.

11. Network and pathway analyses

We analyzed the biological pathways using Ingenuity Pathway Analysis (IPA, ver 23814503, http://www.ingenuity.com/) to evaluate whether these subnetworks are biologically meaningful via comparing the functional relationship between the constituent genes from GWAS. Gene lists from our study and 'Reported Gene' lists from GWAS Catalog (www.genome.gov/gwastudies) were entered into the IPA database. The network score or *P* value represents the significance of the focus gene enrichment and enrichment of focus genes to diseases and functional categories was also evaluated in the IPA Knowledge Base.

12. Statistical Analysis

The association of the selected SNPs with the disease or a subset of the disease was analyzed by comparing minor allele frequencies between case and control groups, with a statistical significance determined by chi-square test, Cohchran-Amitage Trend test (for a codominant model) as parametric methods, and Jonckheere-Terpstra test as a non-parametric method (data not shown) using SAS 9.1.3 version (SAS institute Inc. Cary NC, USA). Also, a logistic regression analysis was used to obtain the odds ratio (OR), 95% confidence interval (CI) for OR, and corresponding *P* values comparing cases with controls regarding the selected SNPs under four alternative models (dominant, recessive, codominant, and allelic). Significant *P* values in association analysis were computed for the replicated samples using the Cochran–Armitage trend test as well as for the combined samples using the Cochran-Mantel-Haenszel test (the Bonferroni correction was applied with statistical significance level $P < 2.0 \times 10^{-3}$ for 25 genes). For minor allele frequency of 0.4, our study achieved 80% power to detect OR as small as 1.56, which was estimated by Quanto (version 1.2.4). For each reported SNP, the power for detection at a nominal *P* value of 0.05 was calculated based on the reported OR and the allele frequency in the Korean population (from the current study).

The prognosis of intestinal BD (the cumulative probabilities of operation, admission, corticosteroid use, and immunosuppressant use after diagnosis) were analyzed using the Kaplan–Meier method, with differences determined

using the log-rank test. Student's t-test and Kruskal-Wallis test were used to compare the differences between sample groups in vitro and in vivo, respectively and p-values < 0.05 were considered statistically significant.

III. RESULTS

1. GWAS and SNP selection for candidate genetic markers of BD

A total of 199 patients with BD (100 BD without intestinal involvement, 99 with intestinal involvement; intestinal BD) and a total of 557 controls with call rates of > 95% in GWAS entered into the case–control analysis. An overview of the clinical characteristics of study samples is provided in Supplementary Table 1. The PCA results suggest a low possibility of false-positive associations resulting from the population admixture (Figure 1A) similar with previous Korean IBD GWAS data.[24, 25] There was a moderate variance between BD cases and controls subjects in our data set on eigenvectors 2, suggesting the genetic heterogeneity arising from the disease (Figure 1B).

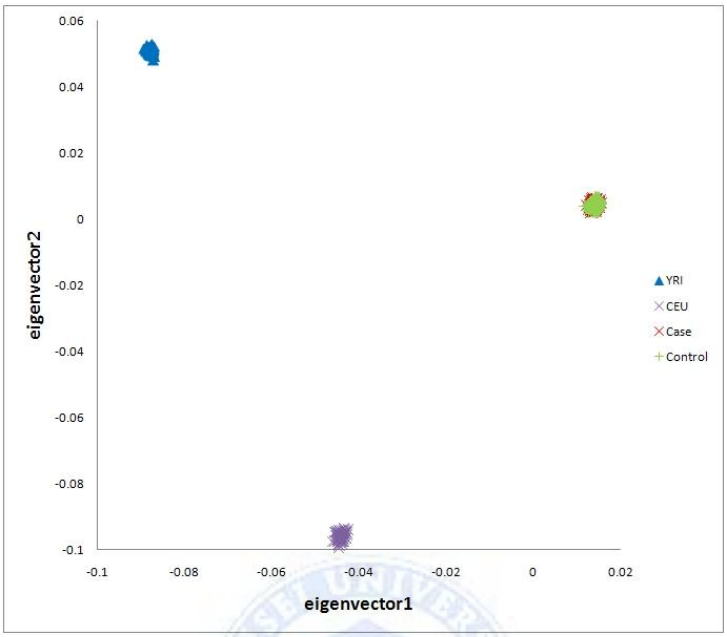
Table 1. Clinical and demographic characteristics of BD without intestinal involvement and intestinal BD patients

	BD without intestinal involvement (n=238)	Intestinal BD (n=295)
Sex, male (%)	62 (26.1)	135 (45.8)
Mean age (years)	41.8±11.8	44.3±12.3
Mean age at diagnosis (years)	NA	40.1±11.8
Mean disease duration (years)	NA	7.5±5.7
HLA-B51 positivity (%)[a]	83 (36.7)	NA
Clinical manifestations (%)		
Oral ulcers	238 (100)	273 (92.5)
Genital ulcers	212 (89.1)	154 (52.2)
Eye lesions	76 (31.9)	52 (17.6)
Skin lesions	219 (92.0)	192 (65.1)
Arthritis/arthralgia	134 (56.3)	158 (53.6)
Vascular lesions	10 (4.2)	11 (3.7)
Central nervous system lesions	5 (2.1)	5 (1.7)
Epididymitis	3 (1.3)	0
Intestinal complications (%)		
Perforation	NA	28 (9.5)
Fistula	NA	26 (8.8)
Stricture	NA	25 (8.5)
Abscess	NA	13 (4.4)
Medication use (%)		
5-aminosalicylic acid/ sulfasalazine	NA	288 (97.6)
Corticosteroids	NA	136 (46.1)
Azathiopurine/6-mercaptopurine	NA	107 (36.3)
Anti-TNF agent	NA	8 (2.7)

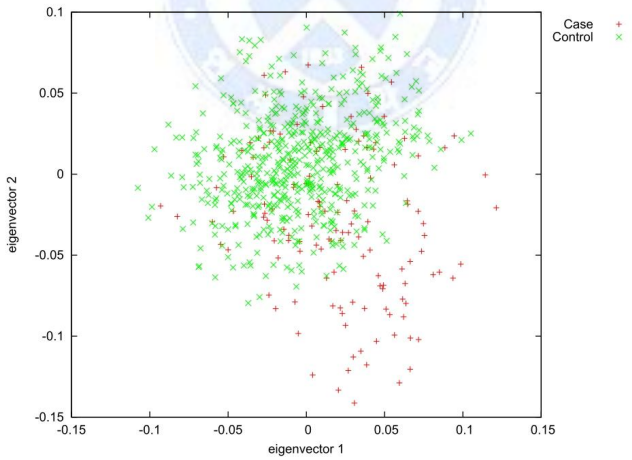
BD, Behcet's disease; HLA, human leukocyte antigen; TNF, tumor necrosis factor; NA, not applicable

[a] Data is available in 226 patients with systemic BD

(A)


0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
eigenvector2
-0.1
-0.08
-0.06
-0.04
-0.02
0
0.02
eigenvector1
YRI
CEU
Case
Control

(B)


0.1
0.05
0
-0.05
-0.1
-0.15
eigenvector 2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
eigenvector 1
Case
Control

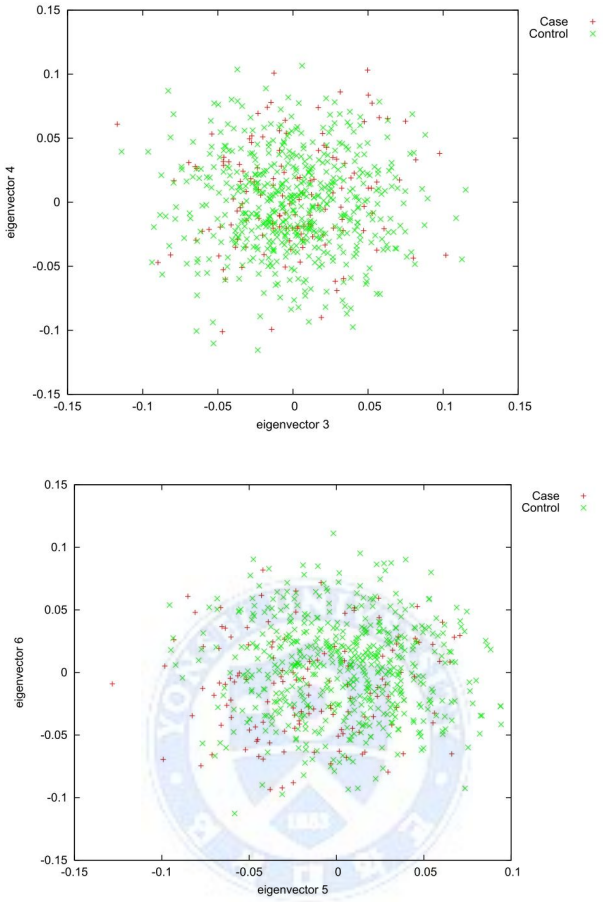


Figure 1. Principal components analysis (PCA). (A) PCA of All participant samples (199 BD case and 597 control) along with reference samples from 1000 Genomes Project. BD cases are shown in red, controls in green, Utah residents with ancestry from northern and western Europe (CEU) in purple, and Yoruba in Ibadan, Nigeria (YRI) in dark blue. (B) PCA of 199 BD case and 597 control samples without samples from 1000 Genomes Project. BD cases are shown in red and controls in green.

After SNP data quality control, we analyzed the association between all BD (intestinal BD+BD without intestinal involvement) and controls for assuming total BD risk, which showed only one significant SNP (rs1681595 in noncoding region) after multiple test correction. Additionally, we tried to analyze subgroups to examine disease phenotypes minutely, for we assumed that intestinal BD is another disease originated from polygenic and differently affected factors different from BD without intestinal involvement. Of these SNPs of $P < 1.0 \times 10^{-4}$, 5 non-clustered SNPs were detected from comparisons between intestinal BD and BD without intestinal involvement among BD-specific 218 SNPs after Bonferroni correction, of which one SNP (rs1681595) with lowest *P* value and 2 genomic loci (rs4500591 in MIR548F5/NBEA and rs32019 in CD180) were subjected to validation test. Moreover, we directly compared intestinal BD with BD without intestinal involvement and 108 significant SNPs were obtained before the multiple test correction, even though we did not detected any significant SNPs after the correction. Of these SNPs, one SNP (rs6838327 in *YIPF7*) with lowest *P* value in clustered genomic locus was subjected to validation test. Additionally, we tried to compare disease groups with healthy control group to explore factors further affecting intestinal BD. Consequently, there were newly identified 6 clustered SNPs including 2 genomic loci (LOC284395, CD300LB) after multiple test correction among significant 342 SNPs including *PSORS1C1* loci in comparison between BD without intestinal

involvement and healthy control. To assess intestinal BD-associated SNPs, we compared intestinal BD with healthy control with 60,137 SNPs not included in regions of BD without intestinal involvement to exclude factors related BD without intestinal involvement, which enable more comprehensive analyzing disease subtypes at the genome-wide level in consideration of polygenic factors that does not acting alone or non-sequencial or independent risk factors. This analysis identified 5 non-clustered SNPs including a genomic locus after multiple test correction. Thus, we selected additionally clustered genomic loci with more than 4 SNP among significant 319 SNPs before multiple test as well as non-clustered SNP with lowest *P* value (rs4500591 in MIR548F5/NBEA).

To improve the coverage of variants specifically associated with intestinal BD, we imputed missing genotypes of autosomal SNPs in our GWAS collection of 99 intestinal BD cases and 557 controls using 1000 Genomes Phase I integrated variant set, from which *CD300C/CD300LB* ($P < 2.0 \times 10^{-6}$) and *TNFRSF13B* ($P < 2.5 \times 10^{-6}$) on chromosome 17 showed significant associations with intestinal BD development.

Following the genome-wide association screening, the 47 selected SNPs were underwent a replication and validation study using additional independent 196 patients with intestinal BD, 137 patients with BD without intestinal involvement, and 391 controls to evaluate the genome-wide association screening. The workflow of this study, Manhattan plots, and

regional association plots that provide a detailed overview of the associated regions is given in Figure 2–4, respectively.

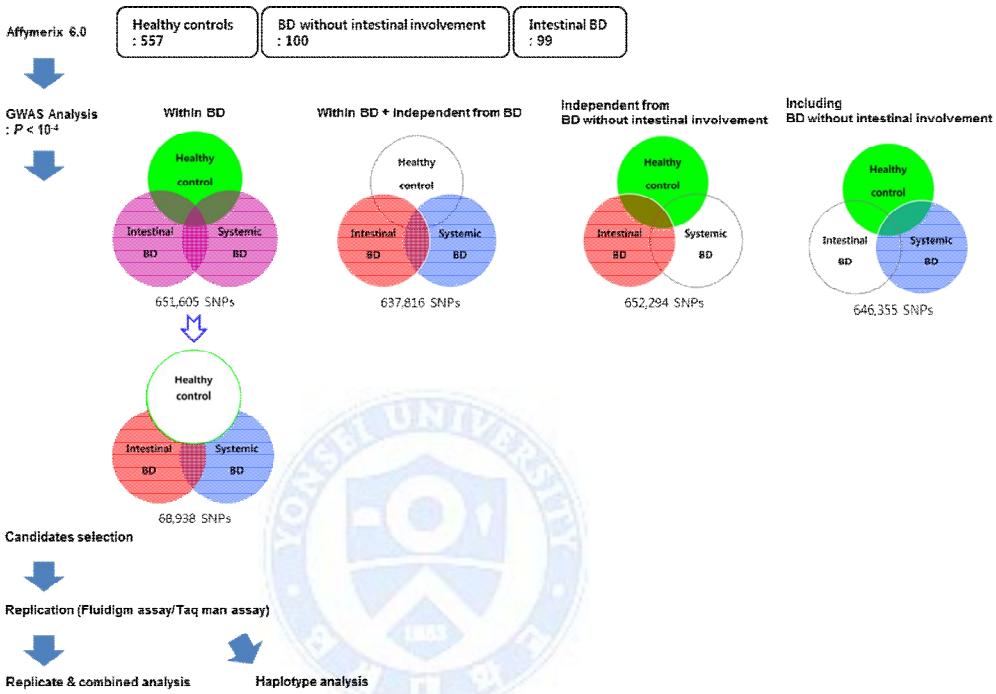
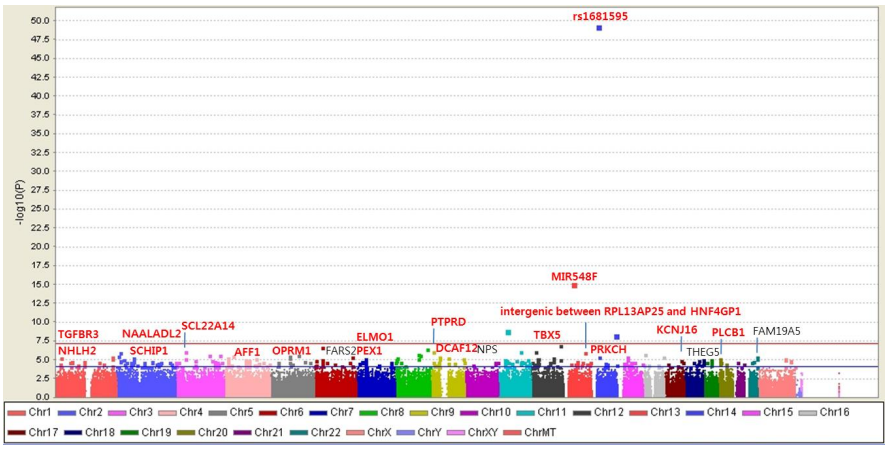
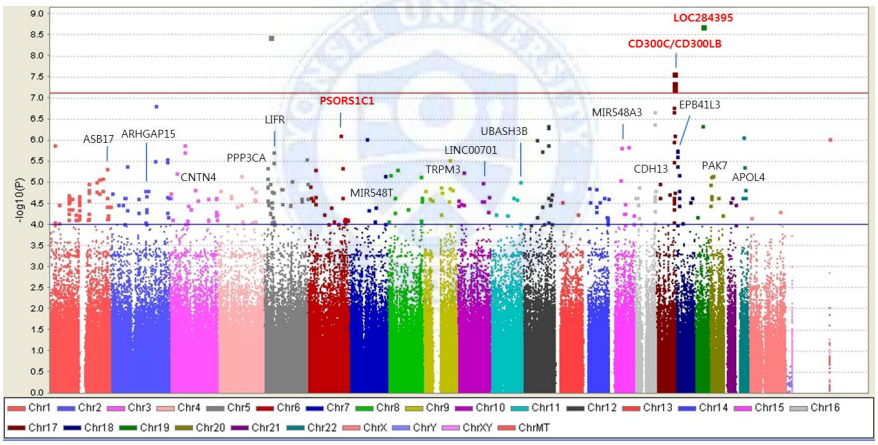


Figure 2. Workflow of the study.

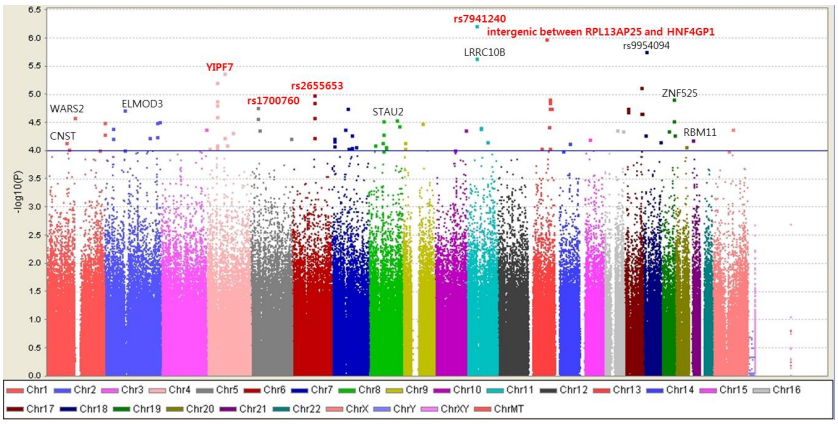
(A)


rs1681595
MIR548F
intergenic between RPL13AP25 and HNF4GP1
TGFBR3
NAALADL2
SCL22A14
NHLH2
SCHIP1
AFF1
OPRM1
FARS2
PEX1
ELMO1
PTPRD
DCAF12
NPS
TBX5
PRKCH
KCNJ16
PLCB1
FAM19A5
THEG5
-log10(P)
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16
Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 ChrX ChrY ChrXY ChrMT

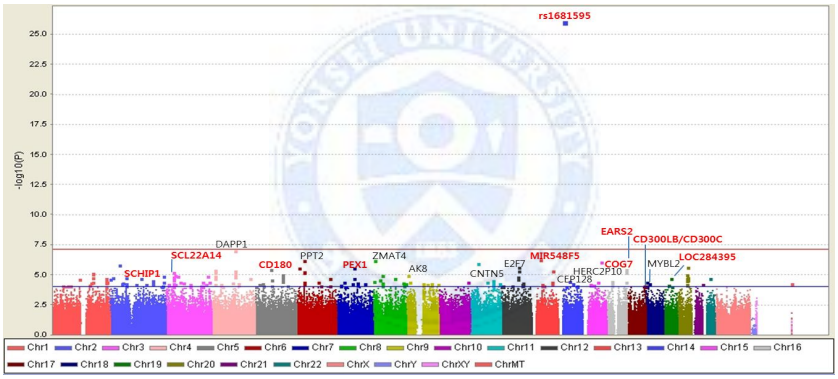
(B)


LOC284395
CD300C/CD300LB
PSORS1C1
MIR548A3
EPB41L3
LIFR
ASB17
ARHGAP15
UBASH3B
LINC00701
PPP3CA
TRPM3
CDH13
PAK7
APOL4
CNTN4
MIR548T
-log10(P)
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16
Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 ChrX ChrY ChrXY ChrMT

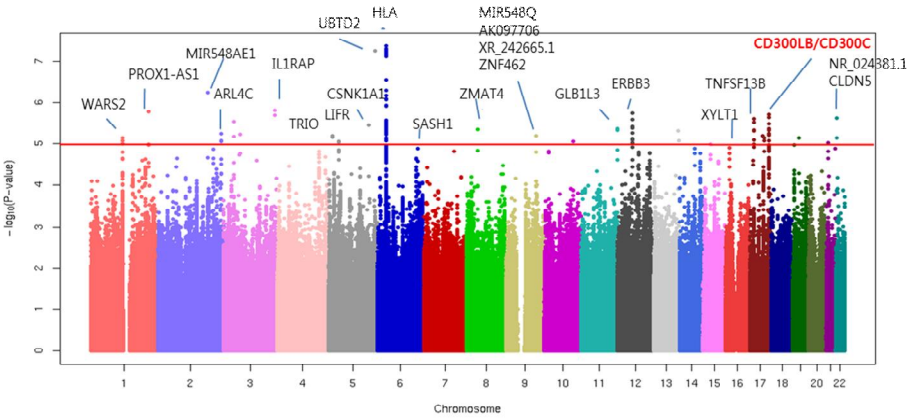
(C)


rs7941240
intergenic between RPL13AP25 and HNF4GP1
rs9954094
LRRC10B
YIPF7
rs2655653
rs1700760
ELMOD3
WARS2
CNST
STAU2
ZNF525
RBM11
-log10(P)
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16
Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 ChrX ChrY ChrXY ChrMT

(D)


rs1681595
EARS2
CD300LB/CD300C
DAPP1
SCL22A14
CD180
PPT2
PEX1
ZMAT4
AK8
MIR548F5
LOC284395
SCHIP1
CNTN5
E2F7
HERC2P10
CEP128
MYBL2
-log10(P)
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12 Chr13 Chr14 Chr15 Chr16
Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 ChrX ChrY ChrXY ChrMT

(E)


USTD2
HLA
MIR548Q
AK097706
XR_242665.1
ZNF462
CD300LB/CD300C
NR_024381.1
CLDN5
MIR548AE1
PROX1-AS1
IL1RAP
ARL4C
WARS2
CSNK1A1
LIFR
TRIO
SASH1
ZMAT4
GLB1L3
ERBB3
TNFSF13B
XYLT1
−log10(P-value)
Chromosome

(F)

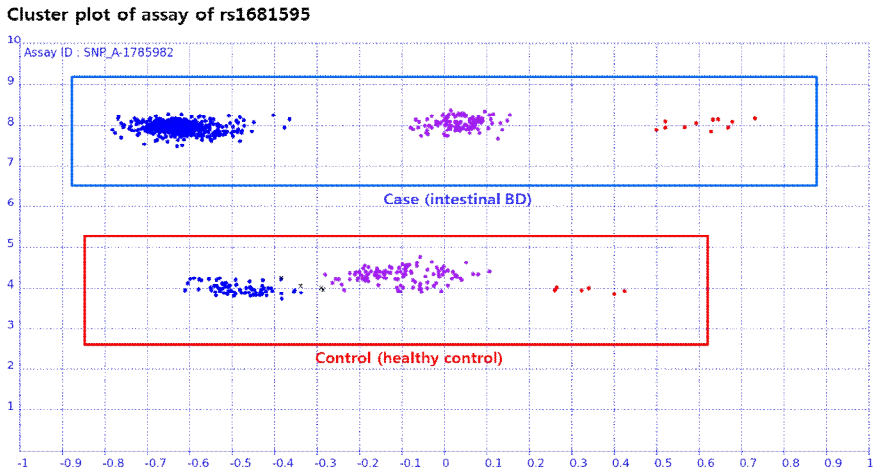
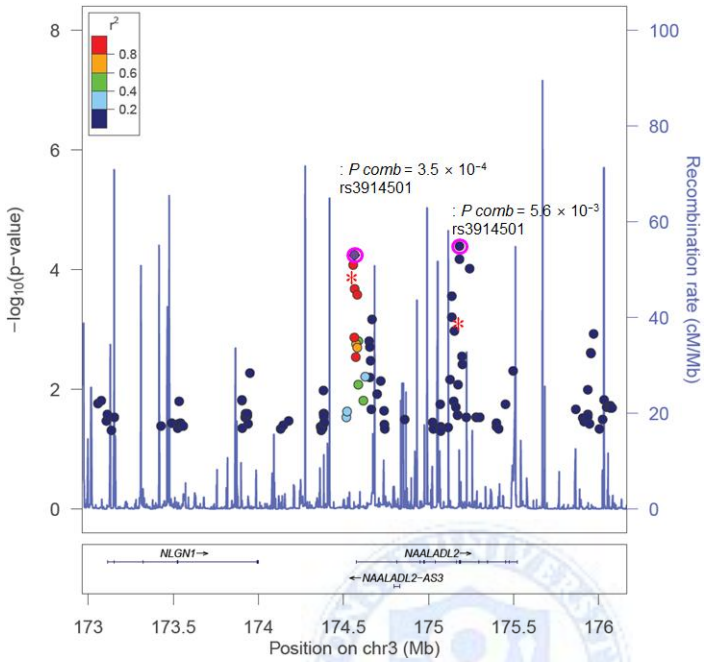
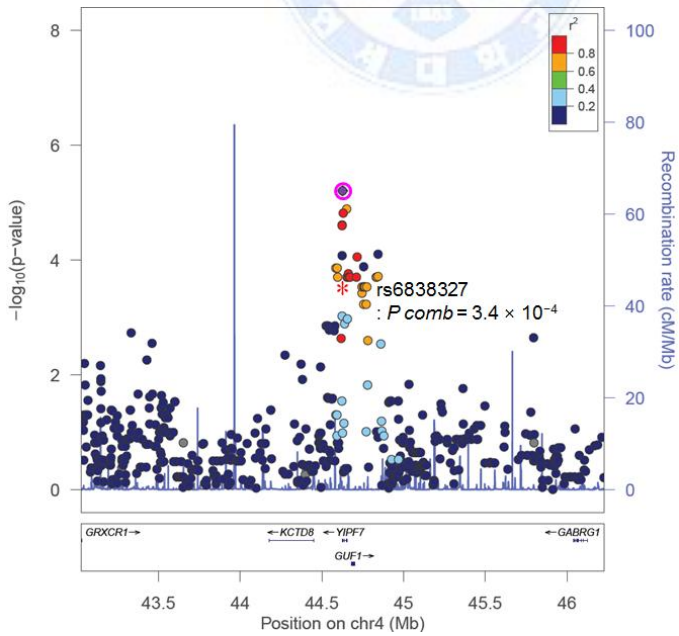


Figure 3. Manhattan plot from comparison between intestinal BD vs. healthy control (A), comparison between BD without intestinal involvement vs. healthy control (B), comparison between intestinal BD vs. BD without intestinal involvement (C), comparison between all BD vs. healthy control (D), and imputated comparison between intestinal BD vs. healthy control (E). (F) Cluster plot of assay of rs1681595 from panel (A). Blue line indicates $P < 10^{-4}$, brown line $P < 10^{-7}$, and red line $P < 10^{-5}$.

(A)


r²
0.8
0.6
0.4
0.2
: P comb = 3.5 × 10⁻⁴
rs3914501
: P comb = 5.6 × 10⁻³
rs3914501
−log10(p−value)
Recombination rate (cM/Mb)
NLGN1
NAALADL2
NAALADL2-AS3
Position on chr3 (Mb)

(B)


r²
0.8
0.6
0.4
0.2
rs6838327
: P comb = 3.4 × 10⁻⁴
−log10(p−value)
Recombination rate (cM/Mb)
GRXCR1
KCTD8
YIPF7
GABRG1
GUF1
Position on chr4 (Mb)

(C)

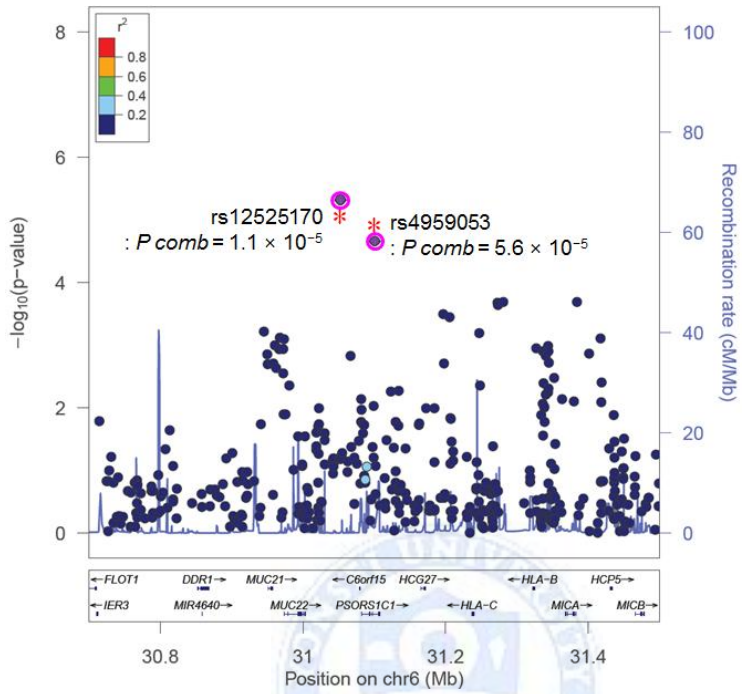


Figure 4. Regional association plots from GWAS ($-\log_{10}$ *P* values; top) and gene loci (bottom). (A) rs3914501 (intestinal BD vs. healthy control), (B) rs6838327 (BD without intestinal involvement vs. intestinal BD), (C) rs4959053 (BD without intestinal involvement vs. healthy control). Single nucleotide polymorphisms (SNPs) are plotted using the estimated recombination rates from hg19/1000 Genomes Asian groups (December 2012). The validated SNPs and the $-\log_{10}$ *P* values were shown by pink circles and red asterisk. Plots were generated using LocusZoom.

2. Replication study for genotype validation for intestinal involvement of BD

The results of 47 selected SNPs from 25 loci showing the strongest association signals or a significant association in more than two regions in at least one of the segmented analyses, 13 SNPs from previously identified 8 gene loci by the BD GWAS studies (rs12119179 and rs1495965 in *IL23R-IL12RB2*; rs1554286, rs1518111, rs1800871 in *IL10*; rs5742906 in *CCR1-CCR3*; rs17482078, rs2927615 in *ERAP1*; rs4713242 in *HLA-F-AS1–HLA-A*; rs2255336 in *KLRC4-KLRK1*; rs12525170, rs4959053 in *PSORS1C1*; rs7574070 in *STAT4*),[2-6, 26] and one locus (rs61730133 in *CD300C/CD300LB*) from our imputed data are selected for the replication study.

Of the selected 43 SNPs, non-clustered SNPs and imputed SNP showed no genetic variation in our study. The comparison between intestinal BD and BD without intestinal involvement showed only one intestinal BD-specific association of *YIPF7* gene locus on chromosome 4 (rs6838327, OR = 1.567, *P* combined = 3.4×10^{-4} in an allelic model) after Bonferroni correction calculated as 0.05/25 tests (Table 2). In addition, a *NAALADL2* locus on chromosome 3 (rs3914501, OR = 1.914, *P* combined = 3.5×10^{-4} in a recessive model) showed significant association between intestinal BD and healthy control (Table 3), and modest evidence of association (rs16848171, OR = 3.208, *P* combined = 5.0×10^{-3} in recessive model before Bonferroni correction) between intestinal BD and BD without intestinal involvement but

no association between BD without intestinal BD and healthy control (Supporting information 4). *PSORS1C1* gene locus (rs12525170) nearby *HLA-B* on chromosome 6 showed a significant association with BD without intestinal involvement (OR = 1.606, $P = 1.1 \times 10^{-5}$ in a dominant model) as well as with BD (OR = 2.210, P combined $= 1.91 \times 10^{-3}$ in a allelic model) after the multiple test correction in both replicated and combined analyses, which showed also a significant association between intestinal BD and BD without intestinal involvement before the multiple test correction (Table 2 and 4). We identified *ELMO1* (rs10259514) as new susceptible locus is associated with BD, which showed a stronger association with intestinal BD (rs10259514, OR = 0.706, $P = 4.8 \times 10^{-3}$ in a codominant model) than BD without intestinal involvement (rs10259514, OR = 0.724, $P = 1.8 \times 10^{-2}$ in codominant model) when compared to healthy controls. This different genetic risk reflects a specific biological significance of the presence of intestinal involvement (Table 2 and 3). Among previously implicated loci in GWAS studies,[2-6, 26] *CCR1-CCR3* (rs5742906), *IL23R-IL12RB2* (rs12119179, rs1495965), *HLA-F-AS1–HLA-A* (rs4713242), and *KLRC4-KLRK1* (rs2255336) and *STAT4* (rs7574070) except *PSORS1C1* (rs4959053) did not show a significant association with any subtypes of BD in our study, whereas *ERAP1* (rs2927615) showed a significant association (OR = 2.744, $P = 1.6 \times 10^{-2}$ in an allelic model) between intestinal BD and BD without intestinal involvement before the multiple test correction, reflecting that *ERAP1*

increases intestinal involvement in BD. In addition, rs1518111 (OR = 0.717, $P = 7.5 \times 10^{-3}$, in a codominant model), rs1554286 (OR = 0.789, $P = 7.3 \times 10^{-3}$ in a codominant model), and rs1800871 (OR = 0.730, $P = 8.9 \times 10^{-3}$ in a codominant model) in *IL10* loci also showed significant association between intestinal BD or BD and healthy control, but no significant association with BD without intestinal involvement. rs7574070 in *STAT4* (OR = 1.355, $P = 4.2 \times 10^{-2}$ in a dominant model) also showed only significant association between BD and healthy control. These results indicate that our data set had an enough power (> 80%) to detect an association for these SNPs, although there is the present study's limitation in sample size. Besides, intestinal BD showed modest evidence of association with *CD180*, *DCAF12*, *PLCB1*, and *TGFBR3* as well as noncoding region *LOC284395*.

Table 2. Association results from comparison between patients with intestinal Behcet's disease and Behcet's disease without intestinal involvement

SNP	MA	Locus	Nearby genes¶	GWAS *P* Model	Replicated a*P* value Model	Combined a*P* value Model	OR (95% CI)†	Combined *P*† Model	MAF	
									BD without ntestinal nvolvement	Intestinal BD
rs3914501	G	Chr. 3: 174564668	NAALADL2	1.1×10^{-1} Allelic	1.1×10^{-1} Allelic	1.1×10^{-1} Allelic	1.303 (0.88–1.928)	1.1×10^{-1} Allelic	).487	0.537
rs16848171	C	Chr. 3: 175181067	NAALADL2	5.0×10^{-3} Recessive	5.0×10^{-3} Recessive	5.0×10^{-3} Recessive	3.208 (1.367–7.529)	5.0×10^{-3} Recessive	).217	0.248
rs32019	C	Chr. 5: 66702373	CD180	9.9×10^{-3} Recessive	9.9×10^{-3} Recessive	9.9×10^{-3} Recessive	1.723 (1.137–2.61)	9.9×10^{-3} Recessive	0.483	0.548

SNP	MA	Locus	Nearby genes[¶]	GWAS *P* Model	Replicated a*P* value Model	Combined a*P* value Model	OR (95% CI)[†]	Combined *P*[†] Model	MAF	
									BD without ntestinal nvolvement	Intestinal BD
rs6086653	A	Chr.20: 8838343	PLCB1	1.5×10^{-2} Recessive	1.5×10^{-2} Recessive	1.5×10^{-2} Recessive	2.141 (1.144–4.007)	1.5×10^{-2} Recessive	).266	0.293
rs284148	T	Chr.1: 92277843	TGFBR3	1.6×10^{-2} Dominant	1.6×10^{-2} Dominant	1.6×10^{-2} Dominant	0.638 (0.442–0.921)	1.6×10^{-2} Dominant	).446	0.381
rs12525170*	A	Chr. 22: 31099761	PSORS1C1	5.8×10^{-3} Dominant	5.8×10^{-3} Dominant	5.8×10^{-3} Dominant	0.595 (0.410–0.862)	5.8×10^{-3} Dominant	).212	0.151
rs121917820	T	Chr. 6: 31084794	PSORS1C1	No variation	No variation	No variation		No variation		
rs7742033	G	Chr. 6: 31085226	PSORS1C1	8.8×10^{-1} Codominant	8.8×10^{-1} Codominant	8.8×10^{-1} Codominant	0.801 (0.050–12.879)	8.8×10^{-1} Codominant	).002	0.002
rs4959053*	A	Chr. 6: 31099577	PSORS1C1	1.1×10^{-2} Dominant	1.1×10^{-2} Dominant	1.1×10^{-2} Dominant	0.619 (0.425–0.9)	1.1×10^{-2} Dominant	).204	0.148
rs5742906*	T	Chr. 3: 46306765	CCR1-CCR3	No variation	No variation	No variation		No variation		
rs10259514	G	Chr. 3: 36829705	ELMO1	5.5×10^{-1} Recessive	5.5×10^{-1} Recessive	5.5×10^{-1} Recessive	1.256 (0.596–2.645)	5.5×10^{-1} Recessive	).256	0.250
rs17482078*	T	Chr. 5: 96118866	ERAP1	1.9×10^{-1} Codominant	1.9×10^{-1} Codominant	1.9×10^{-1} Codominant	1.529 (0.807–2.897)	1.9×10^{-1} Codominant	).032	0.048
rs2927615*	A	Chr. 5: 96198202	ERAP1	1.6×10^{-2} Allelic	1.6×10^{-2} Allelic	1.6×10^{-2} Allelic	2.744 (1.172–6.427)	1.6×10^{-2} Allelic	).016	0.041
rs4713242*	A	Chr. 6: 29718220	HLA-F-AS1–HLA-A	4.0×10^{-1} Recessive	4.0×10^{-1} Recessive	4.0×10^{-1} Recessive	1.227 (0.763–1.971)	4.0×10^{-1} Recessive	).398	0.399
rs1554286*	G	Chr. 1: 206944233	IL10	4.2×10^{-1} Recessive	4.2×10^{-1} Recessive	4.2×10^{-1} Recessive	0.758 (0.381–1.505)	4.2×10^{-1} Recessive	).272	0.252
rs1800871*	C	Chr. 1: 206946634	IL10	4.1×10^{-1} Dominant	4.1×10^{-1} Dominant	4.1×10^{-1} Dominant	0.865 (0.611–1.223)	4.1×10^{-1} Dominant	).266	0.245
rs1518111*	C	Chr. 1: 206944645	IL10	4.1×10^{-1} Allelic	4.1×10^{-1} Allelic	4.1×10^{-1} Allelic	0.889 (0.669–1.18)	4.1×10^{-1} Allelic	0.269	0.247

SNP	MA	Locus	Nearby genes[¶]	GWAS *P* Model	Replicated a*P* value Model	Combined a*P* value Model	OR (95% CI)[†]	Combined *P*[†] Model	MAF	
									BD without ntestinal nvolvement	Intestinal BD
rs12119179[*]	C	Chr. 1: 67747415	IL23R-IL12RB2	4.1×10^{-1} Codominant	4.1×10^{-1} Codominant	4.1×10^{-1} Codominant	0.902 (0.705–1.154)	4.1×10^{-1} Codominant	).515	0.489
rs1495965[*]	C	Chr. 1: 67753508	IL23R-IL12RB2	5.9×10^{-1} Dominant	5.9×10^{-1} Dominant	5.9×10^{-1} Dominant	0.894 (0.593–1.347)	5.9×10^{-1} Dominant	).521	0.514
rs2255336[*]	T	Chr. 12: 10532326	KLRC4-KLRK1	6.6×10^{-1} Recessive	6.6×10^{-1} Recessive	6.6×10^{-1} Recessive	0.801 (0.296–2.169)	6.6×10^{-1} Recessive	).188	0.185
rs7574070[*]	A	Chr. 2: 192010488	STAT4	8.9×10^{-1} Codominant	8.9×10^{-1} Codominant	8.9×10^{-1} Codominant	1.025 (0.685–1.535)	8.9×10^{-1} Codominant	).491	0.487
rs6838327	A	Chr. 4: 44626846	YIPF7	3.4×10^{-4} Allelic	3.4×10^{-4} Allelic	3.4×10^{-4} Allelic	1.567 (1.225–2.005)	3.4×10^{-4} Allelic	).408	0.519
rs7245731	A	Chr. 19: 29975118	LOC284395	1.4×10^{-2} Recessive	1.4×10^{-2} Recessive	1.4×10^{-2} Recessive	0.113 (0.014–0.925)	1.4×10^{-2} Recessive	).139	0.127

The replication study were performed 391 healthy controls and 264 intestinal BD samples including 100 systemic BD samples used in GWAS.[*]Discovered loci described previously. Allele frequencies are presented for the discovery sample. [¶]Nearby genes are defined as the closest genes to the SNP within signal boundary or the closest genes within a 200-kb window. [†]Combined *P*-value, combined OR (CI), or combined MAF obtained from Cochran–Mantel–Haenszel test statistic (1 df). SNP, single nucleotide polymorphism; Chr., chromosome; MA, minor allele; GWAS, genome-wide association study; OR, odds ratio; 95% CI, 95% confidence interval; MAF, minor allele frequency.

Table 3. Association results from comparison between healthy controls and patients with intestinal Behcet's disease

SNP	MA	Locus	Nearby genes¶	GWAS *P* Model	Replicated a*P* value Model	Combined a*P* value Model	OR (95% CI) †	Combined *P*† Model	MAF	
									Healthy control	Intestinal BD
rs3914501	G	Chr. 3: 174564668	NAALADL2	5.7×10^{-5} recessive	1.6×10^{-2} Recessive	3.8×10^{-4} Recessive	1.914 (1.338–2.738)	3.5×10^{-4} Recessive	0.459	0.537
rs16848171	C	Chr. 3: 175181067	NAALADL2	1.4×10^{-5} Recessive	9.2×10^{-2} Recessive	7.0×10^{-3} Recessive	2.462 (1.279–4.739)	5.6×10^{-3} Recessive	0.219	0.248
rs32019	C	Chr. 5: 66702373	CD180	1.8×10^{-3} Codominant	4.4×10^{-2} Codominant	4.0×10^{-3} Codominant	1.265 (1.023–1.563)	3.8×10^{-3} Codominant	0.487	0.548
rs6086653	A	Chr.20: 8838343	PLCB1	3.5×10^{-5} Dominant	6.2×10^{-1} Dominant	4.3×10^{-2} Dominant	0.729 (0.537–0.99)	4.3×10^{-2} Dominant	0.312	0.293
rs284148	T	Chr.1: 92277843	TGFBR3	4.8×10^{-5} Dominant	3.5×10^{-1} Dominant	1.2×10^{-2} Dominant	0.667 (0.486–0.916)	1.2×10^{-2} Dominant	0.425	0.381
rs12525170*	A	Chr. 22: 31099761	PSORS1C1	4.7×10^{-2} Allelic	4.7×10^{-2} Allelic	8.5×10^{-2} Allelic	1.318 (0.962–1.805)	8.5×10^{-2} Allelic	0.119	0.151
rs121917820	T	Chr. 6: 31084794	PSORS1C1	No variation	No variation	No variation		No variation		
rs7742033	G	Chr. 6: 31085226	PSORS1C1	6.4×10^{-1} Codominant	6.4×10^{-1} Codominant	8.4×10^{-1} Codominant	1.323 (0.082–21.241)	8.4×10^{-1} Codominant	0.001	0.002
rs4959053*	A	Chr. 6: 31099577	PSORS1C1	7.6×10^{-2} Allelic	7.7×10^{-1} Allelic	1.2×10^{-1} Allelic	1.285 (0.937–1.760)	1.2×10^{-1} Allelic	0.119	0.148
rs5742906*	T	Chr. 3: 46306765	CCR1-CCR3	No variation	No variation	No variation		No variation		
rs10259514	G	Chr. 3: 36829705	ELMO1	7.7×10^{-6} Allelic	1.4×10^{-1} Allelic	4.8×10^{-3} Allelic	0.706 (0.554–0.899)	4.8×10^{-3} Allelic	0.320	0.249
rs17482078*	T	Chr. 5: 96118866	ERAP1	1.6×10^{-1} Recessive	9.9×10^{-1} Recessive	9.9×10^{-1} Recessive	∞ (0–∞)	2.5×10^{-1} Recessive	0.046	0.048
rs2927615*	A	Chr. 5: 96198202	ERAP1	1.6×10^{-1} Recessive	9.9×10^{-1} Recessive	9.9×10^{-1} Recessive	∞ (0–∞)	2.5×10^{-1} Recessive	0.032	0.041
rs4713242*	A	Chr. 6: 29718220	HLA-F-AS1–HLA-A	9.2×10^{-1} Recessive	9.2×10^{-1} Recessive	5.3×10^{-1} Recessive	1.141 (0.757–1.720)	5.3×10^{-1} Recessive	0.391	0.399
rs1554286*	G	Chr. 1: 206944233	IL10		4.0×10^{-3} Codominant	2.1×10^{-2} Codominant	0.717 (0.562–0.915)	7.3×10^{-3} Codominant	0.318	0.252
rs1800871*	C	Chr. 1: 206946634	IL10	!	5.8×10^{-3} Codominant	2.1×10^{-2} Codominant	0.730 (0.563–0.948)	8.9×10^{-3} Codominant	0.309	0.245
rs1518111*	C	Chr. 1: 206944645	IL10	1.4×10^{-5} Codominant	5.6×10^{-3} Codominant	7.8×10^{-3} Codominant	0.717 (0.562-0.916)	7.5×10^{-3} Codominant	0.313	0.247

SNP	MA	Locus	Nearby genes[¶]	GWAS *P* Model	Replicated a*P* value Model	Combined a*P* value Model	OR (95% CI)[†]	Combined *P*[†] Model	MAF	
									Healthy control	Intestinal BD
rs12119179[*]	C	Chr. 1: 67747415	IL23R-IL12RB2	8.8×10^{-3} Recessive	9.5×10^{-1} Recessive	3.3×10^{-1} Recessive	1.196 (0.832–1.722)	3.3×10^{-1} Recessive	0.478	0.489
rs1495965[*]	C	Chr. 1: 67753508	IL23R-IL12RB2		4.0×10^{-1} Recessive	1.7×10^{-1} Recessive	1.283 (0.901–1.828)	1.7×10^{-1} Recessive	0.494	0.514
rs2255336[*]	T	Chr. 12: 10532326	KLRC4-KLRK1		2.7×10^{-1} Codominant	1.8×10^{-1} Codominant	0.828 (0.630–1.088)	1.8×10^{-1} Codominant	0.214	0.185
rs7574070[*]	A	Chr. 2: 192010488	STAT4		8.0×10^{-2} Dominant	7.3×10^{-2} Dominant	1.370 (0.971–1.934)	7.3×10^{-2} Dominant	0.459	0.491
rs6838327	A	Chr. 4: 44626846	YIPF7		1.2×10^{-1} Recessive	6.4×10^{-3} Recessive	1.654 (1.152–2.376)	6.2×10^{-3} Recessive	0.468	0.519
rs7245731	A	Chr. 19: 29975118	LOC284395		9.9×10^{-1} Recessive	9.9×10^{-1} Recessive	-	2.5×10^{-1} Recessive	0.133	0.127

The replication study were performed 391 healthy controls and 295 intestinal BD samples including 99 intestinal BD samples used in GWAS.[*]Discovered loci described previously. Allele frequencies are presented for the discovery sample. [¶]Nearby genes are defined as the closest genes to the SNP within signal boundary or the closest genes within a 200-kb window. [†]Combined *P*-value, combined OR (CI), or combined MAF obtained from Cochran–Mantel–Haenszel test statistic (1 df). SNP, single nucleotide polymorphism; Chr., chromosome; MA, minor allele; GWAS, genome-wide association study; OR, odds ratio; 95% CI, 95% confidence interval. ap_value: Logistic regression analysis

Table 4. Association results from comparison between healthy controls and Behcet's disease patients without intestinal involvement

SNP	MA	Locus	Nearby genes[¶]	GWAS *P* Model	Replicated a*P* value Model	Combined a*P* value Model	OR (95% CI) [†]	Combined *P*[†] Model	MAF	
									Healthy control	BD without intestinal involvement
rs3914501	G	Chr. 3: 174564668	NAALADL2	8.7×10^{-2} Codominant	7.8×10^{-2} Codominant	3.4×10^{-1} Codominant	1.122 (0.886–1.422)	3.4×10^{-1} Codominant	0.459	0.487
rs16848171	C	Chr. 3: 175181067	NAALADL2	6.2×10^{-1} Recessive	4.1×10^{-1} Recessive	5.7×10^{-1} Recessive	0.768 (0.308–1.911)	5.7×10^{-1} Recessive	0.219	0.217
rs32019	C	Chr. 5: 66702373	CD180	2.4×10^{-1} Recessive	3.3×10^{-1} Recessive	3.9×10^{-2} Recessive	0.658 (0.442–0.98)	3.9×10^{-2} Recessive	0.487	0.483
rs6086653	A	Chr.20: 8838343	PLCB1	2.8×10^{-1} Codominant	1.9×10^{-1} Codominant	7.9×10^{-2} Codominant	0.792 (0.61–1.028)	7.9×10^{-2} Codominant	0.312	0.266
rs284148	T	Chr.1: 92277843	TGFBR3	4.4×10^{-1} Recessive	8.4×10^{-1} Recessive	3.0×10^{-1} Recessive	1.25 (0.823–1.897)	3.0×10^{-1} Recessive	0.425	0.446
rs12525170[*]	A	Chr. 6: 31207740	PSORS1C1		1.1×10^{-3} Dominant	1.1×10^{-5} Dominant	2.210 (1.546–3.157)	1.1×10^{-5} Dominant	0.119	0.212
rs121917820	T	Chr. 6: 31084794	PSORS1C1		No variation	No variation		No variation		
rs7742033	G	Chr. 6: 31085226	PSORS1C1		2.0×10^{-1} Codominant	7.2×10^{-1} Codominant	1.648 (0.103–26.477)	7.2×10^{-1} Codominant	0.001	0.002
rs4959053[*]	A	Chr. 6: 31099577	PSORS1C1	2.2×10^{-5} Allelic	7.7×10^{-2} Allelic	5.6×10^{-5} Allelic	1.892 (1.383–2.589)	5.6×10^{-5} Allelic	0.119	0.203
rs5742906[*]	T	Chr. 3: 46306765	CCR1-CCR3	No variation	No variation	No variation		No variation		
rs10259514	G	Chr. 3: 36829705	ELMO1	6.8×10^{-2} Codominant	1.7×10^{-2} Codominant	1.7×10^{-2} Codominant	0.724 (0.556–0.944)	1.7×10^{-2} Codominant	0.320	0.256
rs17482078[*]	T	Chr. 5: 96118866	ERAP1		5.3×10^{-2} Codominant	2.1×10^{-1} Codominant	0.669 (0.358–1.251)	2.1×10^{-1} Codominant	0.046	0.032
rs2927615[*]	A	Chr. 5: 96198202	ERAP1		5.0×10^{-2} Codominant	7.5×10^{-2} Codominant	0.468 (0.199–1.100)	7.5×10^{-2} Codominant	0.032	0.016

SNP	MA	Locus	Nearby genes[¶]	GWAS *P* Model	Replicated a*P* value Model	Combined a*P* value Model	OR (95% CI)[†]	Combined *P*[†] Model	MAF	
									Healthy control	BD without intestinal involvement
rs4713242[*]	A	Chr. 6: 29718220	HLA-F-AS1–HLA-A		6.4×10^{-1} Dominant	5.7×10^{-1} Dominant	1.104 (0.788–1.548)	5.7×10^{-1} Dominant	0.391	0.397
rs1554286[*]	G	Chr. 1: 206944233	IL10	1.4×10^{-1} Codominant	5.7×10^{-1} Codominant	8.5×10^{-2} Codominant	0.801 (0.621–1.032)	8.5×10^{-2} Codominant	0.318	0.272
rs1800871[*]	C	Chr. 1: 206946634	IL10		4.6×10^{-3} Dominant	1.0×10^{-1} Dominant	0.806 (0.622–1.044)	1.0×10^{-1} Dominant	0.309	0.266
rs1518111[*]	C	Chr. 1: 206944645	IL10		3.2×10^{-3} Dominant	1.0×10^{-1} Dominant	0.758 (0.545–1.055)	1.0×10^{-1} Dominant	0.313	0.278
rs12119179[*]	C	Chr. 1: 67747415	IL23R-IL12RB2	1.1×10^{-1} Recessive	2.2×10^{-1} Recessive	1.0×10^{-1} Recessive	1.369 (0.938–2.000)	1.0×10^{-1} Recessive	0.478	0.515
rs1495965[*]	C	Chr. 1: 67753508	IL23R-IL12RB2		8.2×10^{-1} Recessive	2.0×10^{-1} Recessive	1.276 (0.877–1.855)	2.0×10^{-1} Recessive	0.494	0.523
rs2255336[*]	T	Chr. 12: 10532326	KLRC4-KLRK1	3.3×10^{-1} Allelic	1.2×10^{-1} Allelic	2.7×10^{-1} Allelic	0.85 (0.637–1.134)	2.7×10^{-1} Allelic	0.214	0.188
rs7574070[*]	A	Chr. 2: 192010488	STAT4	4.0×10^{-1} Dominant	8.7×10^{-2} Dominant	1.2×10^{-1} Dominant	1.337 (0.924–1.933)	1.2×10^{-1} Dominant	0.459	0.487
rs6838327	A	Chr. 4: 44626846	YIPF7	3.9×10^{-4} Dominant	3.6×10^{-1} Dominant	1.0×10^{-2} Dominant	0.630 (0.442–0.897)	1.0×10^{-2} Dominant	0.468	0.408
rs7245731	A	Chr. 19: 29975118	LOC284395	1.9×10^{-9} Recessive	9.9×10^{-1} Recessive	6.3×10^{-4} Recessive	-	6.3×10^{-4} Recessive	0.133	0.139

The replication study were performed 391 healthy controls and 264 intestinal BD samples including 100 systemic BD samples used in GWAS.[*]Discovered loci described previously. Allele frequencies are presented for the discovery sample. [¶]Nearby genes are defined as the closest genes to the SNP within signal boundary or the closest genes within a 200-kb window. [†]Combined *P*-value, combined OR (CI), or combined MAF obtained from Cochran–Mantel–Haenszel test statistic (1 df). SNP, single nucleotide polymorphism;

Chr., chromosome; MA, minor allele; GWAS, genome-wide association study; OR, odds ratio; 95% CI, 95% confidence interval; MAF, minor allele frequency.

3. Haplotype analysis

For multifactorial diseases, single variant on a particular gene may have no strong effect, but the combination of multiple variants with small effects explains the overall susceptibility to the disease.[27] To better evaluate genetic associations at haplotype level, we selected 11 gene loci from GWAS data and conducted a haplotype analysis with the additional proxy SNPs in the *DCAF12*, *ERAP1*, *IL10*, *IL23R-IL12RB2*, *SCHIP1*, *NAALADL2*, *PLCB1*, *PSORS1C1*, and *TGFBR3* showing multiple signals from GWAS. While *IL23R-IL12RB2* did not show an association with any type of BD, *DCAF12, IL10*, *NAALADL2*, *PLCB1*, *SCHIP1*, and *TGFBR3* showed associations with intestinal BD development in haplotype analysis, which was consistent with our replication results: *DCAF12* (C-A, OR = 0.619, $P = 2.4 \times 10^{-3}$ between intestinal BD vs. healthy control), *IL10* (G-C-C, OR = 0.712, $P = 7.3 \times 10^{-3}$ between intestinal BD vs. healthy control), *NAALADL2* (T-G, OR = 0.473, $P = 1.7 \times 10^{-2}$ between intestinal BD vs. healthy controls; T-G, OR = 0.347, $P = 7.3 \times 10^{-3}$ between intestinal BD vs. BD without intestinal involvement), *PLCB1* (C-C-T-T-G, OR = 1.465, $P = 1.5 \times 10^{-2}$ between intestinal BD vs. healthy control), *SCHIP1* (G-C, OR = 1.613, $P = 3.2 \times 10^{-3}$ between intestinal

BD vs. healthy control), and *TGFBR3* (C-C-G, OR = 1.481, $P = 2.2 \times 10^{-2}$ between intestinal BD vs. healthy control), suggesting these genes to be potential causal variants contributing to intestinal BD development (Table 5). Among these loci, the significant associations of *NAALADL2* and *PSORS1C1* were remained in comparisons between intestinal BD and BD without intestinal involvement. On the other hands, only *PSORS1C1* showed a significant association with BD without intestinal involvement (G-G-G, OR = 0.511, $P = 1.5 \times 10^{-5}$ between BD without intestinal involvement vs. healthy control), which might have a negative correlation for intestinal involvement.

Table 5. Halplotype analysis of susceptible genes for intestinal Behcet disease and Behcet disease without intestinal involvement

Closest genes	Haplotype	Combined $P^{\dagger}$	Model	OR (95% CI) †
Intestinal BD (n = 295) vs. healthy control (n = 391)				
DCAF12	C-A	2.4×10^{-3}	Dominant	0.619 (0.454–0.843)
	T-G	5.4×10^{-3}	Recessive	1.55 (1.138–2.111)
IL10	A-T-T	1.2×10^{-2}	Codominant	1.374 (1.073–1.76)
	G-C-C	7.3×10^{-3}	Codominant	0.712 (0.555–0.913)
NAALADL2	C-G	5.5×10^{-2}	Recessive	2.079 (0.985-4.388)
	T-C	6.4×10^{-1}	Dominant	1.137 (0.66-1.961)
	T-G	1.7×10^{-2}	Dominant	0.473 (0.255–0.876)
PLCB1	C-C-T-T-G	1.5×10^{-2}	Recessive	1.465 (1.077–1.993)
	G-T-C-C-A	9.1×10^{-2}	Dominant	0.768 (0.565-1.044)
SCHIP1	C-T	7.3×10^{-3}	Recessive	0.649 (0.473–0.890)

Closest genes	Haplotype	Combined $P^{\dagger}$ Model	OR (95% CI) †
	G-C	3.2×10^{-3} Dominant	1.613 (1.174–2.217)
TGFBR3	C-C-G	2.2×10^{-2} Recessive	1.481 (1.059–2.070)
	C-T-A	1.4×10^{-1} Dominant	0.793 (0.582-1.080)
BD without intestinal involvement (n =264) vs. healthy controls (n = 391)			
PSORS1C1	G-A-A	2.4×10^{-5} Allelic	1.935 (1.424–2.628)
	G-G-G	1.5×10^{-5} Allelic	0.511 (0.377–0.693)
	A-A	3.7×10^{-5} Allelic	1.917 (1.412–2.602)
	G-G	2.1×10^{-5} Allelic	0.515 (0.380–0.699)
Intestinal BD (n = 295) vs. BD without intestinal involvement (n =264)			
NAALADL2	C-G	1.7×10^{-2} Recessive	3.401 (1.245–9.296)
	T-C	4.1×10^{-1} Dominant	0.791 (0.453-1.382)
	T-G	7.3×10^{-3} Dominant	0.347 (0.16–0.751)
PSORS1C1	G-A-A	9.7×10^{-3} Dominant	0.617 (0.428–0.889)
	G-G-G	9.4×10^{-3} Recessive	1.619 (1.125–2.329)
	A-A	9.7×10^{-3} Dominant	0.617 (0.428–0.889)
	G-G	9.6×10^{-3} Recessive	1.62 (1.125–2.333)

SNPs were excluded if they had a minor allele frequency (MAF) < 0.05 and SNPs distant > 100 kB. The risk allele is underlined. †Combined *P*-value, combined OR (CI), or combined MAF obtained from Cochran–Mantel–Haenszel test statistic (1 df). OR, odds ratio; 95% CI, 95% confidence interval; MAF, minor allele frequency.

4. Clinical phenotypes and outcomes according to genotypes

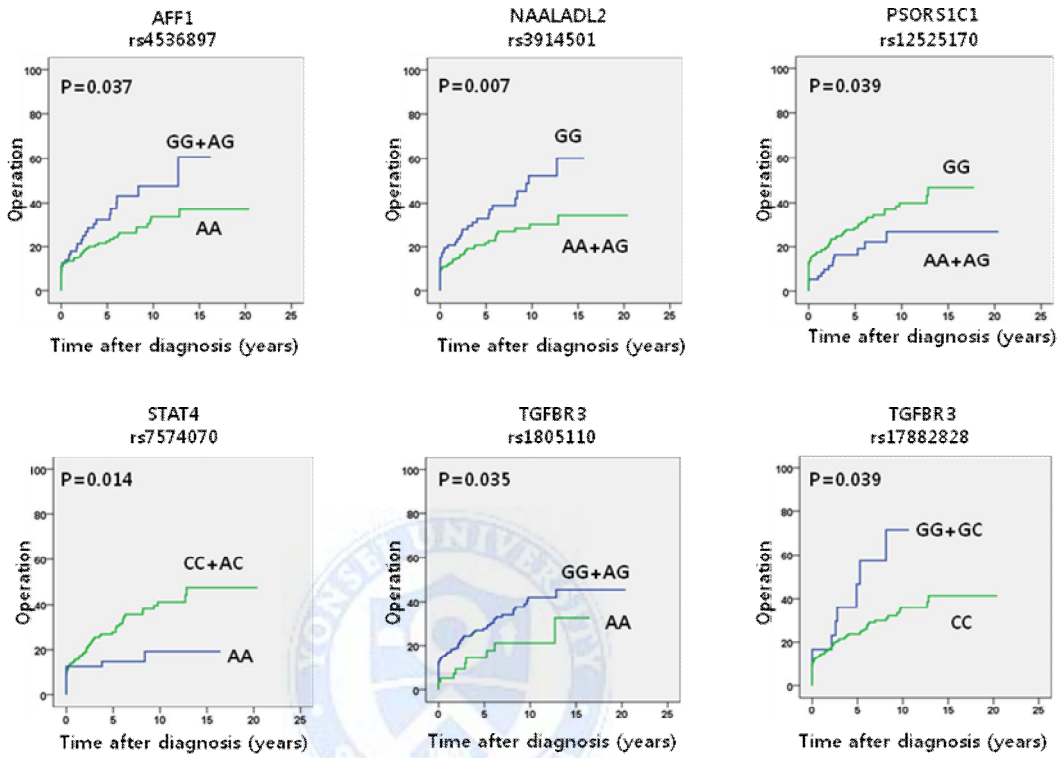
We assessed whether the investigated SNPs could have influence clinical manifestations including intestinal complications and prognosis such as

cumulative probabilities of operation, hospitalization, corticosteroid use, and immunosuppressant use in intestinal BD patients. In intestinal BD, *ERAP1* (rs17482078 and rs2927615) and *HLA-F-AS1–HLA-A* (rs4713242) were associated with bowel perforation. *ELMO1* (rs10259514) and *DCAF12* (rs10758242) were associated with intestinal fistula development, whereas *TGFBR3* (rs1805110 and rs284148) was associated with intestinal stricture (Table 6). In terms of clinical outcomes, *NAALADL2* (rs3914501) and *TGFBR3* (rs1805110 and rs17882828) risk alleles were associated with a higher cumulative probability of surgery, whereas *PSORS1C1* (rs12525170) and *STAT4* (rs7574070) risk alleles were associated with a lower cumulative probability of surgery. *NAALADL2* (rs3914501 and rs16848171) risk alleles were associated with a higher cumulative probability of hospitalization. *NAALADL2* (rs3914501 and rs16848171) and *IL23R-IL12RB2* (rs1495965) risk alleles were associated with a higher cumulative probability of corticosteroid use, suggesting more poor prognosis (Figure 5).

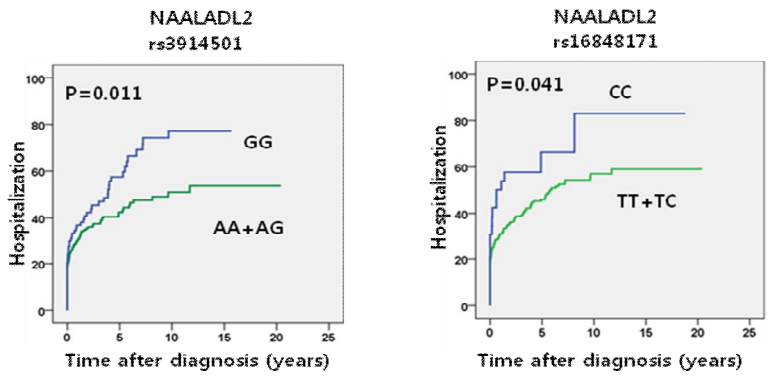
Table 6. Intestinal complications according to SNP genotypes in intestinal BD patients

Phenotype	Gene and rs number		Genotype or allele	Patients with pheneotype (%)	Patients without pheneotype (%)	aOR (95% CI)	*P* value
Perforation	ERAP1 rs17482078	dominant	CC CT/TT	20(76.9) 6 (23.1)	244(92.1) 21 (7.9)	3.49 (1.26-9.62)	0.016
	ERAP1 rs2927615	dominant	GG GA/AA	20 (76.9) 6 (23.1)	246 (93.5) 17 (6.5)	4.34 (1.54-12.24)	0.005
	HLA-F-AS1–HLA-A rs4713242	allele	A G	14 (26.9) 38 (73.1)	218 (41.3) 310 (58.7)	0.52 (0.28-0.99)	0.047
Fistula	ELMO1 rs10259514	allele	G A	19(39.58) 29(60.42)	123(23.65) 397(76.35)	2.12 (1.15-3.90)	0.017
	DCAF12 rs10758242	dominant	GG AG/AA	19(79.2) 5 (20.8)	142(55.7) 113 (44.3)	0.33 (0.12-0.91)	0.033
Stricture	TGFBR3 rs1805110	dominant	GG GA/AA	4 (16.0) 21 (84.0)	101 (37.7) 167 (62.3)	3.17 (1.06-9.51)	0.039
	TGFBR3 rs284148	allele	T C	26 (59.1) 18 (40.9)	192 (36.5) 334 (63.5)	2.51 (1.34-4.70)	0.004

(A)


AFF1
rs4536897
P=0.037
GG+AG
AA
Operation
Time after diagnosis (years)
NAALADL2
rs3914501
P=0.007
GG
AA+AG
Operation
Time after diagnosis (years)
PSORS1C1
rs12525170
P=0.039
GG
AA+AG
Operation
Time after diagnosis (years)
STAT4
rs7574070
P=0.014
CC+AC
AA
Operation
Time after diagnosis (years)
TGFBR3
rs1805110
P=0.035
GG+AG
AA
Operation
Time after diagnosis (years)
TGFBR3
rs17882828
P=0.039
GG+GC
CC
Operation
Time after diagnosis (years)

(B)


NAALADL2
rs3914501
P=0.011
GG
AA+AG
Hospitalization
Time after diagnosis (years)
NAALADL2
rs16848171
P=0.041
CC
TT+TC
Hospitalization
Time after diagnosis (years)

(C)

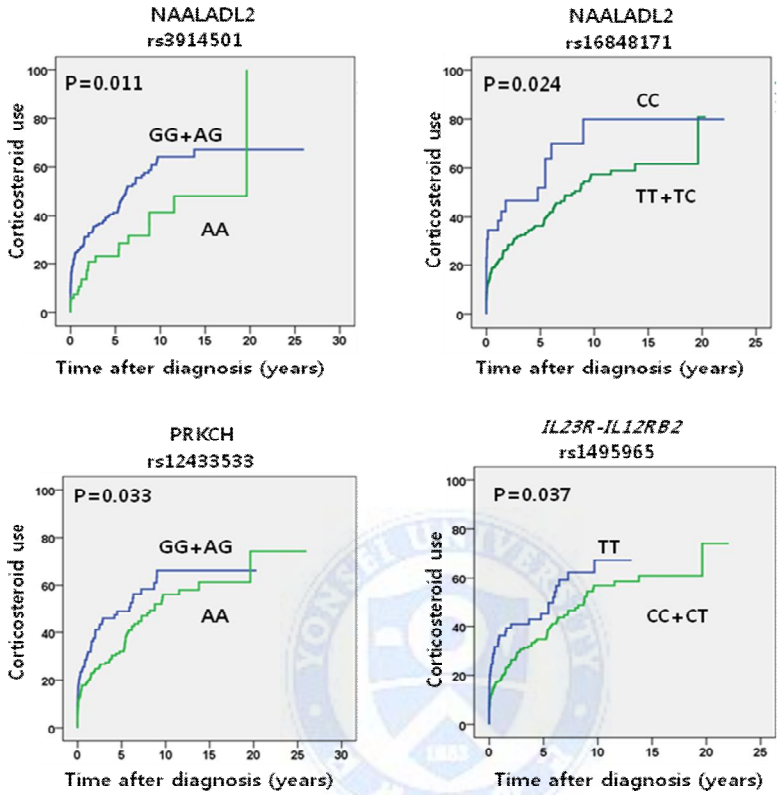


Figure 5. Cumulative probabilities of surgery (A), hospitalization (B), and corticosteroid use (C) according to SNP genotypes in intestinal BD patients. Blue line show risk allele.

We also evaluated whether the investigated SNPs could influence clinical manifestations in BD without intestinal involvement. *PSORS1C1* (rs12525170 and rs4959053) and *IL10* (rs1554286) were related with *HLA-B51* positivity, while *YIPF7* (rs6838327) was related with development of genital ulcers.

IL10 (rs1554286 and rs1957895) was associated with skin involvement and *HLA-F-AS1–HLA-A* (rs4713242) was associated with vascular involvement. *PSORS1C1* (rs12525170 and rs4959053) was related with central nervous system involvement (Table 7).

Table 7. Clinical manifestations according to SNP genotypes in BD patients without intestinal involvement

Phenotype	Gene and rs number		Genotype or allele	Patients with pheneotype (%)	Patients without pheneotype (%)	aOR (95% CI)	P value
HLA-B51 positivity	PSORS1C1 rs12525170	dominant	GG AG/AA	11 (13.6) 70 (86.4)	127 (90.1) 14 (9.9)	56.39 (24.22-131.32)	<0.001
	PSORS1C1 rs4959053	dominant	GG AG/AA	11 (11.9) 68 (86.1)	128 (90.8) 13 (9.2)	59.53 (25.26-140.30)	<0.001
	IL10 rs1554286	dominant	AA GA/GG	37 (44.6) 46 (55.4)	81(57.0) 61 (43.0)	1.75 (1.00-3.05)	0.048
Genital ulcers	YIPF7 rs6838327	allele	A G	177 (42.6) 239 (57.5)	13 (56.0) 37 (74.0)	2.11 (1.08-4.11)	0.029
Skin lesions	IL10 rs1554286	codominant	AA GA GG	120 (54.8) 84 (38.4) 15 (6.9)	6 (33.3) 9 (50.0) 3 (16.7)	0.47 (0.23-0.99)	0.049
	IL10 rs1957895	dominant	TT GT/GG	79(36.1) 140 (63.9)	13(72.2) 5 (27.8)	4.82 (1.63-14.23)	0.004
CNS lesions	PSORS1C1 rs12525170	recessive	GG/AG AA	3 (60.0) 2 (40.0)	221 (96.5) 8 (3.5)	23.23 (3.05-176.79)	0.002
	PSORS1C1 rs4959053	recessive	GG/AG AA	3 (60.0) 2 (40.0)	218 (96.5) 8 (3.5)	22.91 (3.02-174.09)	0.002

5. Gene function analysis

To establish biological relevance those SNPs to disease, we explored the possible functional consequences of the selected 47 SNPs, focusing on existing regulatory SNP and QTL databases containing various cell lines and tissues. rs17482078 (p.Arg725Gln) in *ERAP1* and rs9866564 (p.Pro622Arg) in *NAALADL2* causing missense mutations were predicted to possibly damage in silico analysis. For further investigation, we performed immunohistochemical analysis with colon tissues from intestinal BD patients and normal tissues from cancer patients after intestinal surgery for NAALADL2 and YIPF7 which were significantly associated with intestinal BD development by more than one method. We detected expression of NAALADL2 and YIPF7 in the intestinal epithelium, in which protein expressions of NAALADL2 and YIPF7 were associated with increased proportions of abnormal crypts and significantly decreased in the colon tissues of patients with IBD and BD when compared to control tissues (Figure 6).

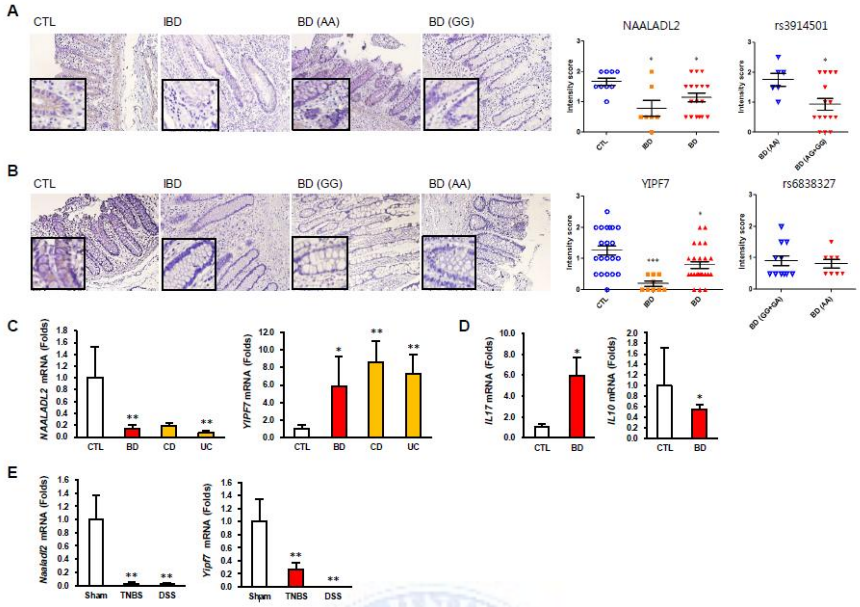


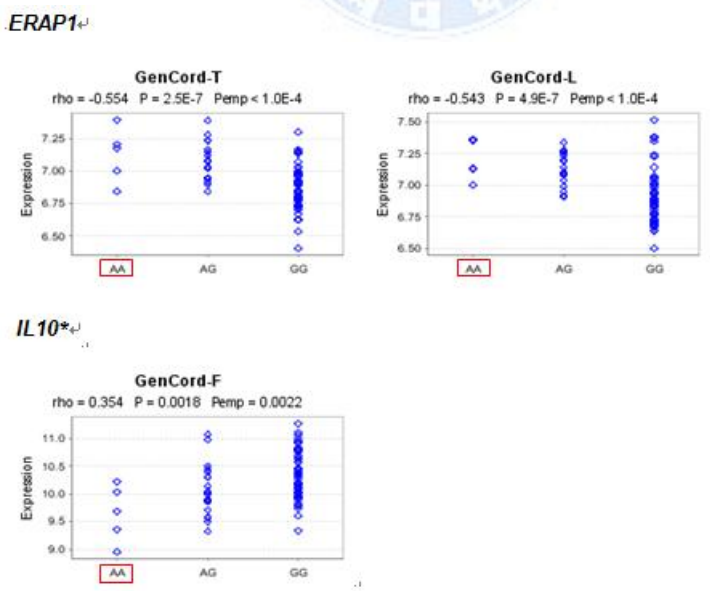
Figure 6. Gene expression in human and mouse colon tissues. (A, B) Immunohistochemistry of human colon tissues colon tissue was performed against NAALADL2 (A) and YIPF7 (B) and the expression of according to diseases and SNP genotypes (NAALADL2: rs3914501; YIPF7: rs6838327) was evaluated. Right panels show stain intensity. Sections were counterstained with hematoxylin. (C–E) mRNA levels in colon tissues of human and mouse. Total RNAs were isolated and converted into cDNA from mouse colon tissues. Transcript levels of *NAALADL2* (C, left), *YIPF7* (C, right), *IL17* (D, left), and *IL10* (D, right) from human colon tissues and *Naaladl2* (E, left), *Yipf7* (E, right) from mouse were quantified by real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) and normalized to *Glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) or *β-*

actin. * $P < 0.05$ vs. CTL, ** $P < 0.005$ vs. CTL or Sham. BD, Behcet diseases; CD, Crohn's disease; CTL, control; UC, ulcerative colitis.

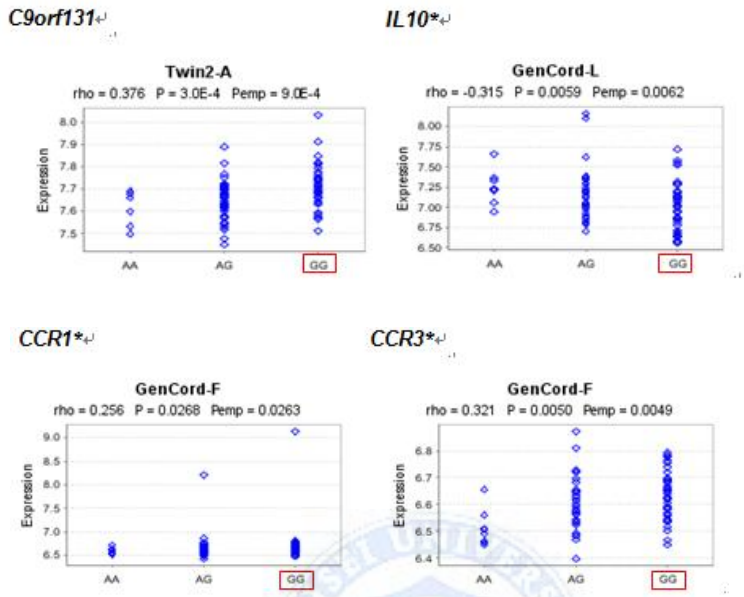
Furthermore, the candidate SNPs were examined for potential regulatory functions using RegulomeDB annotating with reference to experimentally supported regulatory elements. Two of the 3 SNPs indicating a relatively high degree of evidence for potential regulatory function were located in *IL10* (rs1554286, score =1f; rs1518111, score = 3a) and in *TGFBR3* (rs17882828, score= 2b), which are affecting the binding of CREB Binding Protein (CREBBP) and CCAAT/Enhancer Binding Protein beta (CEBPB) involved in colon inflammation[28] and IL17 signaling pathway.[29] To explore additional regulatory mechanisms for the linked SNP set on haplotype blocks using LD information, we use the HaploReg tool. The search of HaploReg indicated that rs3914501 and rs16848171 in *NAALADL2*, rs12525170 in *PSORS1C1*, rs10259514 in *ELMO1*, rs10758242 in *DCAF12*, rs1805110 and rs284148 in *TGFBR3* were located in predicted regulatory motifs (Cdx, TAL1; Alx4, GR, Smad4; TBX5; Hltf, YY1; AIRE, AP-3, Alx4; Osr; PLAG1), suggesting a potential regulatory role. In the high LD block with rs6086653 in *PLCB1*, there are SNPs that may affect regulatory motifs (rs6086632 and rs6039302: Foxp1; rs12624809: E2F, Irf; rs6086633: NERF1a, Pou2f2, TEF; rs6086653: Esr2, Spz1). Furthermore, rs17482078 in *ERAP1* and rs1805110 and rs284148 in *TGFBR3* located in an evolutionary conserved region predicted to be under

functional constraint based on SiPhy analysis. The eQTL analyses revealed a significant associated locus (rs2927615) with its own expressions (*ERAP1*) after Bonferroni corrections. In addition, risk alleles of rs2927615 including rs10441723 (*DCAF12*), rs1805110 (*ELMO1*), rs1554286 (IL10) showed down-regulation of *IL10* expression in lymphocytes and fibroblasts. Moreover, risk alleles of rs10441723 and rs1805110 in *ELMO1*, rs1554286 in *IL10*, rs1805110 in *TGFBR3*, rs7245731 in *LOC284395*, and rs4959053 in *PSORS1C1*, showed the gene expressions of pro-inflammatory cytokines, *CCR1*, *CCR3*, and *IL12RB2* as well as *KLRC4*, *PSORS1C1*, and *STAT4* (Figure 7). These results suggest that many of the risk variants have multiple regulatory and functional features at its locus.

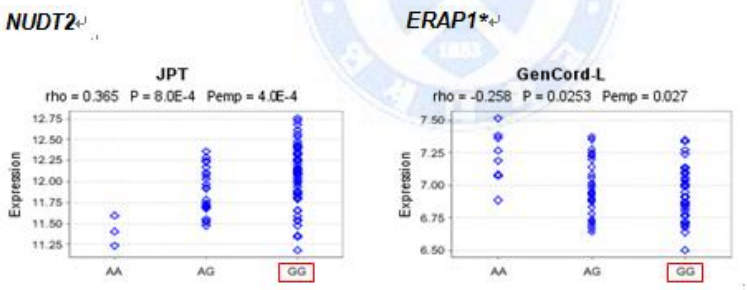
(A) rs2927615 in *ERAP1*



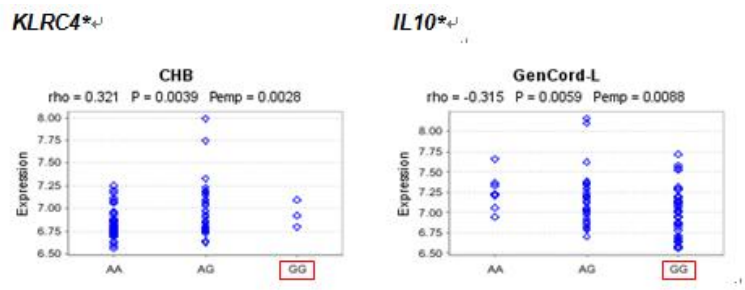
(B) rs10441723 in *DCAF12*

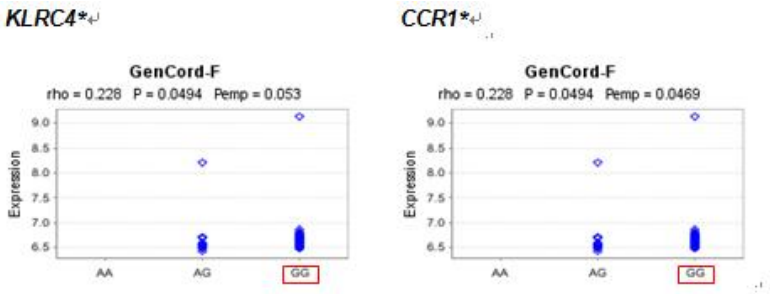


(C) rs10758242 in *DCAF12*

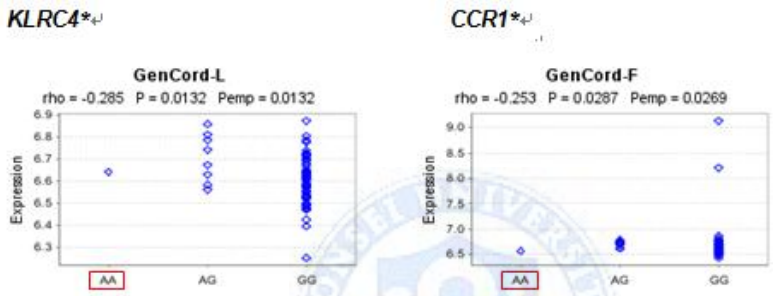


(D) rs1805110 in *ELMO1*

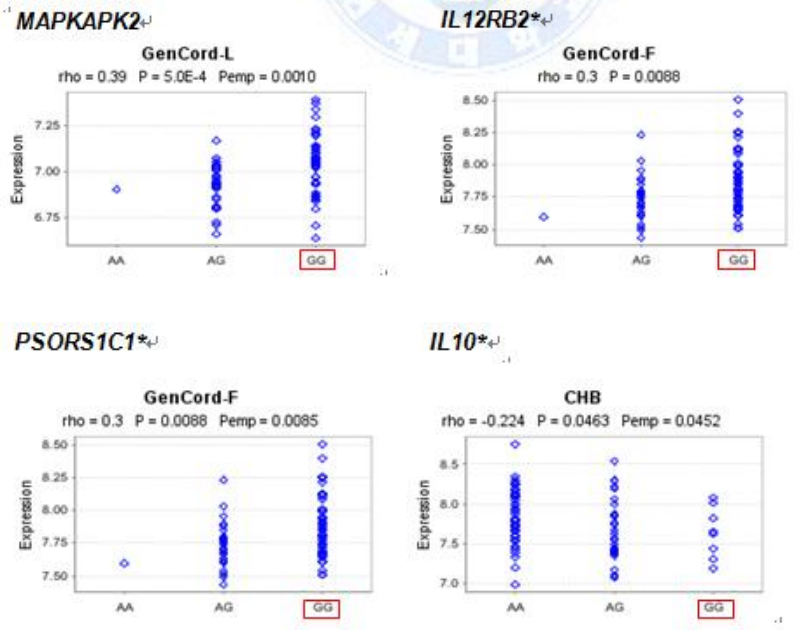




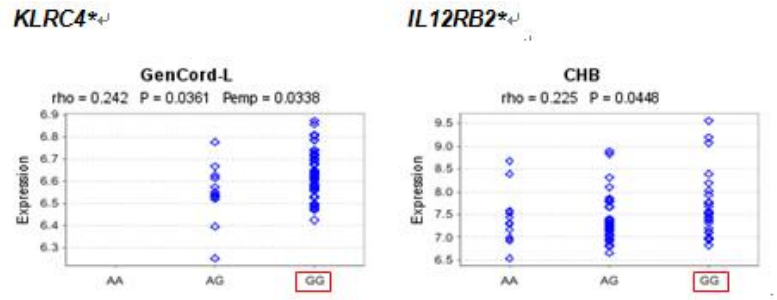
(E) rs4959053 in *PSORS1C1*



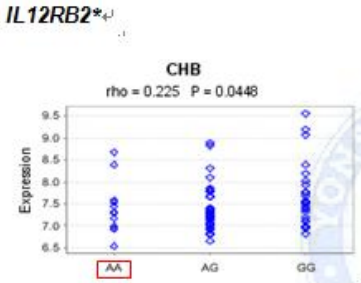
(F) rs1554286 in *IL10*



(G) rs1805110 in *TGFBR3*



(H) rs6838327 in *YIPF7*



(I) rs7245731 in *LOC284395*

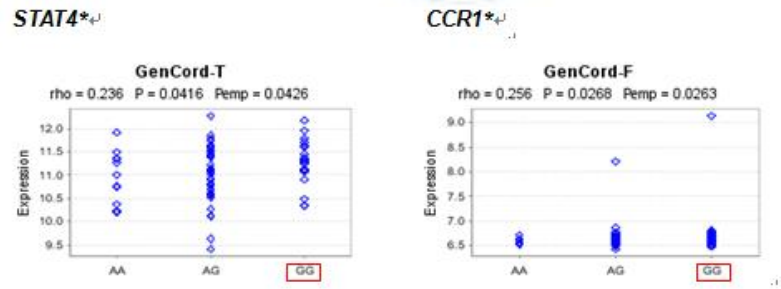


Figure 7. Expression quantitative trait loci (eQTL) analysis. (A) rs2927615 in *ERAP1*. (B) rs10441723 in *DCAF12*. (C) rs10758242 in *DCAF12*. (D) rs1805110 in *ELMO1*. (E) rs4959053 in *PSORS1C1*. (F) rs1554286 in *IL10*.

(G) rs1805110 in *TGFBR3*. (H) rs6838327 in *YIPF7*. (I) rs7245731 in *LOC284395*. Results were plotted from probes located of each SNP of HapMap3 (CEU, CHB, GIH, JPT, LWK, MEX, MKK, YRI; Stranger et al., 2012) and Geneva Gencord (Dimas et al., 2009) studies. A: adipose tissue; F, fibroblast; T, T-cell; S, skin; F, fat; L, lymphoblastoid cell line. Asterisks and red boxes show a significant eQTLs among candidate SNPs for BD after 10,000 permutations and risk alleles, resepctively.

To identify the potential biological pathways with the genes responsible for disease susceptibility from GWAS and the replicated SNPs, we used IPA based on published data. We examined overlap of our significant regions against published GWAS results (GWAS catalog). One functional network of 3 identified networks from the lists (14 genes) from BD GWAS catalog included 9 focus molecules, which has a potential function in T helper cell differentiation ($P = 2.83 \times 10^{-8}$), IL12 signaling and production in macrophage ($P = 3.79 \times 10^{-7}$), IL10 signaling ($P = 4.88 \times 10^{-4}$), and dendritic cell maturation ($P = 3.31 \times 10^{-3}$). The top functional network of 25 networks from lists (536 genes) of IBD GWAS catalog included 26 focus molecules and has a potential function in T helper cell differentiation ($P = 2.97 \times 10^{-29}$), dendritic cell maturation ($P = 8.27 \times 10^{-24}$), and crosstalk between dendritic cells and natural killer cells ($P = 5.7 \times 10^{-22}$).

Notably, the pathways analyses showed that the intestinal BD genes (13 genes) have one functional network included 11 focus molecules and overlapped with 8 networks (Network 5, 9, 10, 13, 15, 16, 22, 23) of IBD GWAS genes, suggesting that the phenotype of intestinal involvement shares common risk factors with IBD (Figure 8*A* and Figure 9*B*). Although intestinal BD genes have only one network (Network 1) directly connected with genes from BD GWAS, these interactions may also connect indirectly intestinal BD pathway with IBD pathway (networks 2, 11: Figure 8*B*). In addition, the functional network shares common potential function in PI3K signaling in B lymphocytes (PLCB1, CD180) and CXCR4 (PLCB1, ELMO1) signaling with 2 focus molecules (CD180, ELMO1) from the lists (5 genes) from BD without intestinal involvement (Figure 9*A*), suggesting that intestinal BD genes shows BD phenotypes through interaction with genes with BD without intestinal involvement as well as intestinal involvement. In addition, comparisons between BD genes and IBD genes from GWAS catalog, 2 networks of BD (Network 1, Network 3) were overlapped with IBD networks, in which Network 1 is linked to Network 2 and Network 11 of IBD networks (Figure 9*C*). These results support the potential for crosstalk between intestinal BD and BD without intestinal involvement.

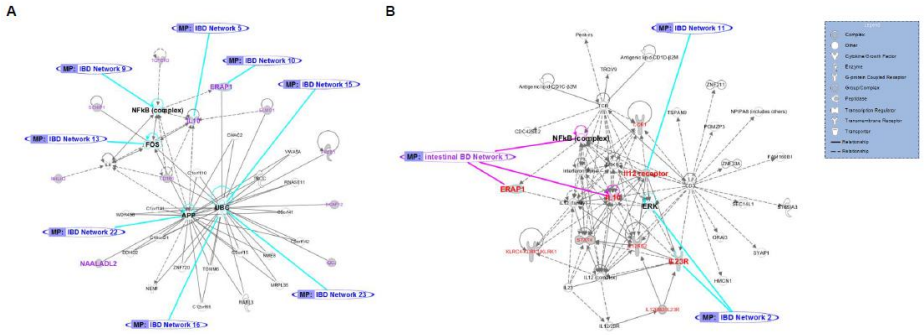
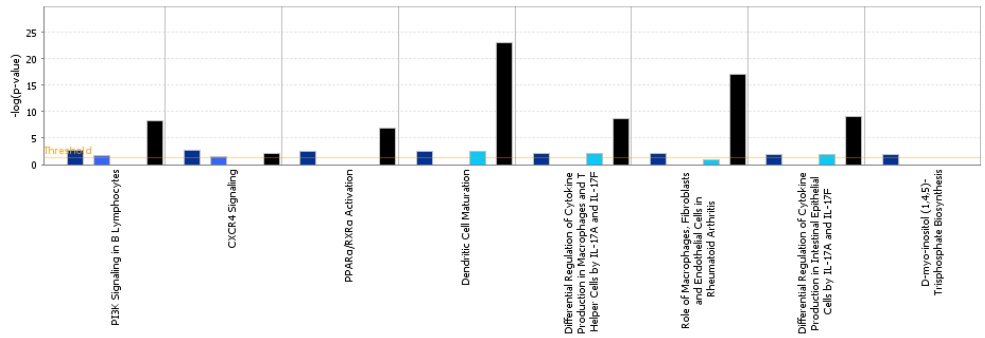
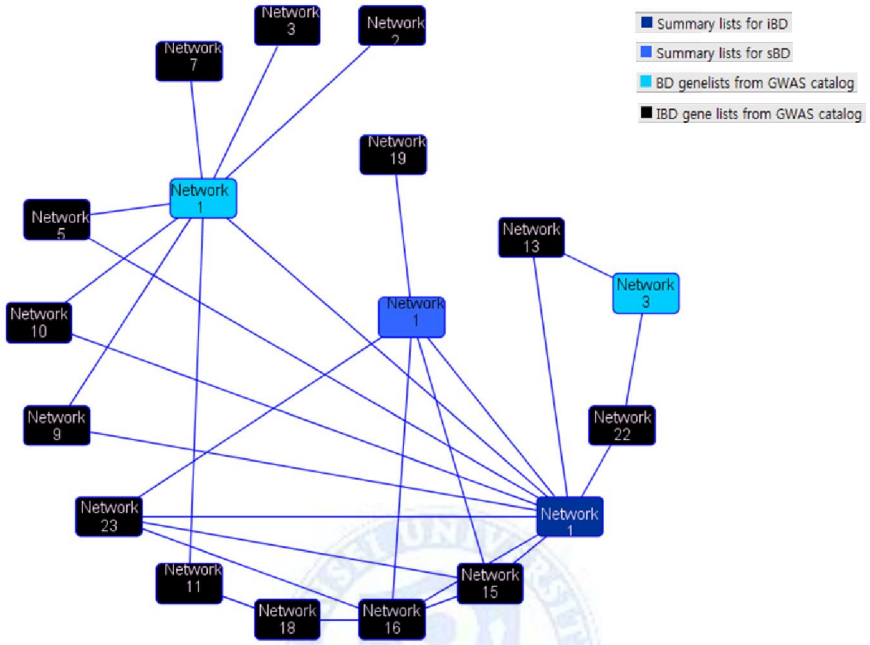


Figure 8. Pathway analyses of genes susceptible to intestinal BD, BD without intestinal involvement, and IBD. (A) Networks of overlapping between intestinal BD and IBD. (B) Connection of intestinal involvement phenotype of BD through networks between BD without intestinal involvement and IBD. Genes implicated by the SNPs shown to be associated with intestinal BD and BD without intestinal involvement at genome-wide significance are highlighted in purple and red, respectively. A core analysis was run on these genes with default settings.

(A)



(B)


Summary lists for iBD
Summary lists for sBD
BD genelists from GWAS catalog
IBD gene lists from GWAS catalog
Network 3
Network 2
Network 7
Network 19
Network 1
Network 5
Network 13
Network 3
Network 10
Network 1
Network 9
Network 22
Network 23
Network 1
Network 11
Network 15
Network 18
Network 16

(C)

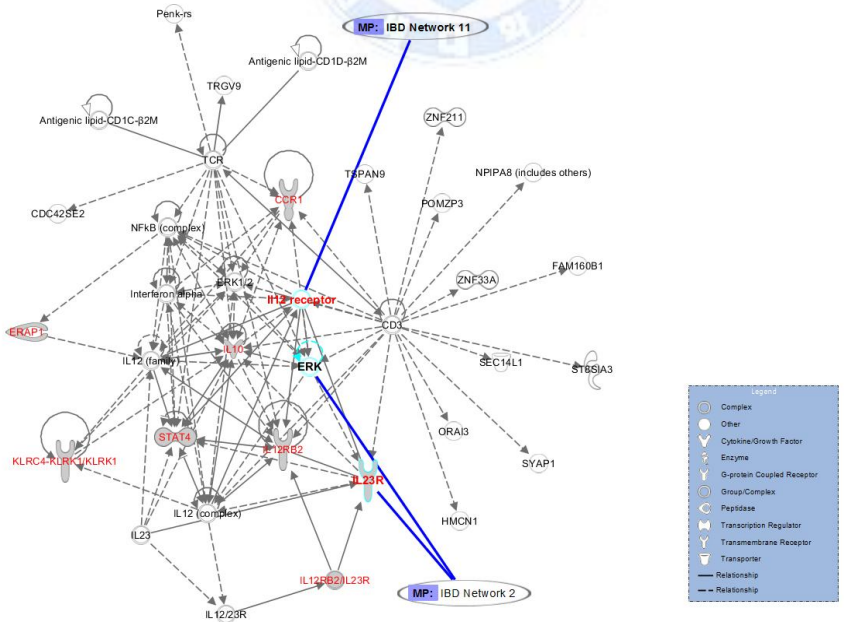

MP: IBD Network 11
Penk-rs
Antigenic lipid-CD1D-β2M
TRGV9
Antigenic lipid-CD1C-β2M
TCR
ZNF211
TSPAN9
NPIPA8 (includes others)
CDC42SE2
CCR1
POMZP3
NFkB (complex)
FAM160B1
ZNF33A
Interferon alpha
ERK1/2
IL12 receptor
CD3
ERAP1
IL10
IL12 (family)
ERK
SEC14L1
ST8SIA3
ORAI3
STAT4
IL12RB2
SYAP1
KLRC4-KLRK1/KLRK1
IL23R
IL12 (complex)
HMCN1
IL23
IL12RB2/IL23R
MP: IBD Network 2
IL12/23R
Legend
Complex
Other
Cytokine/Growth Factor
Enzyme
G-protein Coupled Receptor
Group/Complex
Peptidase
Transcription Regulator
Transmembrane Receptor
Transporter
Relationship
Relationship

Figure 9. Network Diagram of BD susceptible genes. (A) Comparisons of pathways of intestinal BD, BD without intestinal involvement, and IBD. (B) Networks of overlapping among intestinal BD, BD without intestinal involvement, and IBD. (C) Networks of overlapping between BD without intestinal involvement and IBD. A core analysis was run on these genes with default settings. Genes implicated by the SNPs shown to be associated with BD without intestinal involvement at genome-wide significance are highlighted in red.

IV. DISCUSSION

This is the first study to elucidate the specific genetic polymorphisms contributing to the susceptibility to intestinal involvement in BD. Comparisons within intestinal BD and healthy control or BD without intestinal involvement showed intestinal BD-specific associations of locus nearby *NAALADL2* (N-acetyl-L-aspartyl-L-glutamate peptidase I) gene and *YIPF7* (Yip1 domain family member 7) gene. In addition, a computational protein prediction suggested that the missense mutations of *NAALADL2* and *YIPF7* are possibly damaging to their corresponding proteins. Moreover, *NAALADL2*-associated rs3914501 risk allele was related with higher cumulative probabilities of surgery, hospitalization, and corticosteroid use. Our results suggest that *NAALADL2* gene may be associated with poor

prognosis as well as development of intestinal BD. Previously, the rs17531088, *NAALADL2*-associated risk allele, was reported to be linked to development of Kawasaki disease and Cornelia de Lange syndrome, a rare developmental malformation syndrome.[30, 31] Of note, SNP in *NAALADL2* (rs62285902) was also observed in CD genome-wide imputation analysis from Japanese population as candidate gene (OR = 0.63 (0.62–0.65)), although the SNP was not validated in replication study.[32] Considering that Kawasaki disease affects the blood vessels and BD is also a form of vasculitis, it is interesting that *NAALADL2* showed the novel association with intestinal BD in our replication analysis as well as subsequent haplotype analysis replication study. Dysfunction of YIP1 family member may dysregulate intestinal homeostasis leading to pathogenic states.[33] In addition, mice in which *Yipf6* functioning in membrane trafficking and vesicle biogenesis was null mutated were extremely sensitive to colitis induced by dextran sodium sulfate and regarded as a susceptibility locus of IBD in humans in one study. We identified more frequent variations of *YIPF7*, known as another YIP1 family member including Yipf6, in intestinal BD compared to BD without intestinal involvement. *YIPF7* was not included in the pathway analysis results, probably because the exact function of this protein is still unknown. Our eQTL analysis demonstrated that rs3914501 in *NAALADL2* increase the expression of IL13 and IL22 and rs6838327 in *YIPF7* increase the expression of *IL12RB2* in lymphoblastoid cells.

IL-23 drives expression of IL-17 in CD4+ T cells coupled with the distinct functional roles of IL-12,[34] in which IL-22 and CEBPB was induced and act in concert to mediate the pro-inflammatory effects.[35] CEBPs are known to be critical regulators of IL6 expression.[35] Our data show increased IL17 expression in colon tissues of BD patients along with decreased IL10 expression. Patients with CD exhibited increased lamina propria IL12 production.[36]

Furthermore, the IHC showed that the expression of NAALADL2 and YIPF7 not only was significantly decreased in the colonic enterocytes from intestinal BD and IBD patients compared to control but also more decreased in those of BD with risk alleles than in those of BD without risk alleles. Indeed, these results were supported in inflamed colon by colitis in mice and LPS stimulation of intestinal epithelial cells reduced the expression of NAALADL2 and YIPF7 in vitro. SNPs in *NAALADL2* contain characteristics of transcriptional regulatory activity directly impacting its expression. Concordantly, LPS reduced also the mRNA level of *NAALADL2* both in vivo and in vitro. However, mRNA levels of YIPF7 were higher in colon of intestinal BD and IBD than in colon of control, while lower in inflamed colon tissues of mice than in colon tissues of control. It is possible that YIPF7 may be regulated post-transcriptional and translational regulation or caused by different pathophysiology of intestinal BD and IBD compared with animal model using chemical.

Taken together, our results suggest that these genes are strong candidates for the causal variants of intestinal BD development.[37, 38] Interestingly, *HLA-B*51* was associated with moderately higher prevalence of genital ulcers, ocular and skin manifestations, and a decreased prevalence of gastrointestinal involvement in BD[39] in GWAS, whereas *HLA-B*51* showed no relationship with IBD.[40] Consistently, we found that *PSORS1C1* gene loci in the HLA-B region showed significant association with BD without intestinal involvement through GWAS, replication, and haplotype studies. Moreover, *PSORS1C1* is most significant associated loci with all BD consistent with previous BD study in Han Chinese in GWAS study,[3, 6] whereas *HLA-B* did not in our study. Although *HLA-B*51* is known as the most strongly associated genetic risk factor to BD in several populations,[3, 4, 41] it accounts for less than 20% of the whole genetic risk.[41] In addition, the independent genetic associations of *HLA-B*51*, *PSORS1C1*, and *HLA-F-AS1* with BD reported in a previous study for fine mapping in the extended HLA region were consistent with our data[3]. Moreover, *PSORS1C1* showed strongly association with HLA-B51 positivity (OR: 56.39) and CNS lesions (aOR: 23.23), which is stronger interaction compared to interaction (OR: 3.78) between *ERAP1* and *HLA-B*51* found in BD GWAS.[2] The possible role of HLA-B*51 in BD may be the presentation of HLA-B*51-restricted peptides to $CD8^+$ cells and the in inhibitory effect of HLA-B*51 binding to killer Ig-like receptor DL1 (KIR3DL1) on NK cells,[42, 43] although the exact clinical pictures of HLA-

B*51-positive and negative BD are virtually indistinguishable and genotyping of this allele cannot accurately predict the occurrence of specific organ or system manifestations[39] Taken together, our results suggest that *PSORS1C1* genetic variation is associated with development of BD without intestinal involvement independently and regardless of *HLA-B* as previous study[3] and made the strongest contribution to BD without intestinal involvement in our study.

Among previously implicated loci of GWAS,[4, 5] *CCR1-CCR3*, *IL23R-IL12RB2*, *HLA-F-AS1–HLA-A*, and *KLRC4-KLRK1* did not show a significant association with any subtypes of BD in our study. One of the main reasons might be that individual has a unique combination of factors involved in their disease development, GWAS does not take account of supporting evidence from more-complex patterns of associations with other polymorphisms in the same gene and the relatively modest sample size compromises power[44] as well as other unknown variants in the region in high LD with the investigated loci may be altering the regulation of the protein expression, or influencing protein translation and degradation.[45, 46] Hence, the analysis of functional anlysis and pathways may help further elucidate genetic factors in BD pathogenesis. *IL10* and *STAT4* polymorphisms were associated with BD development with > 80% power before Bonferroni correction, which corresponds to other recent studies,[5, 47] indicating that the conditions of our data set have statistically enough power. In addition, *ERAP1* and *ELMO1* showed modest evidence in

terms of susceptibility of intestinal BD and *ERAP1* and *ELMO1* was associated with intestinal perforation and intestinal fistula, respectively, which suggest that these genes may be related with severe disease phenotypes. Consistently, previsous studies implicated that ERAP1 are shared in inflammatory pathways in BD and spondyloarthritides[48] and SNPs in *ERAP1* showed association with Japanese CD patients.[32] ELMO1 (engulfment and cell motility 1) promotes internalization of the dying cells through the ability to engulf and to clear apoptotic cells[49]. ERAP1 is a ubiquitous metallopeptidase that functions to trim cooperatively with ERAP2 antigenic peptide precursors for proper loading onto MHC class I molecules and *ERAP1*-deficient mice were less able to present self- and foreign antigens.[50-52] rs17482078 and rs2927615 in *ERAP1* showed possibly damaging in silico analysis and the regulatory function in eQTL analysis, respectively. Consistent with these results, non-synonymous coding variant rs17482078 (p.Arg725Gln) confer protection against these diseases through reduced peptide trimming and antigen presentation by MHC class I molecules[48] and recessively conferred the risk of disease with uveitis.[2] eQTL analyses showed that rs2927615 in *ERAP1*, located in an evolutionary conserved sequence, not only down-regulates its own gene expression including *STAT4* and *PSORS1C1* but also up-regulates *IL10* expression in company with SNPS in *ELMO1* and *DCAF12*. Regulation of IL10 expression is of particular interest because it protects colonic inflammation in mice[53-56] and as well as other

inflammatory diseases.[57] SNPs of *IL10* were shown to have more significant correlation with intestinal BD than BD without intestinal involvement in our replication and haplotype studies and correlation with down-regulation of its expression and increase of proinflammatory gene expression such as *IL12RB2* and *PSORS1C1* in eQTL analysis. The risk loci of intestinal BD (*IL10*) were shared with IBD,[58-60] the other loci of BD without intestinal involvement (*PSORS1C1*) with psoriasis,[61, 62] which implicates distinguishing and shared pathogenic pathways between BD and other autoimmune diseases as well as between intestinal BD and IBD. As support these results, the 13 replicated loci related to intestinal BD function in B lymphocytes and CXCR4, which is shared with BD without intestinal involvement, where the common variations in each of these genes have little correlation with disease risk. In addition, pathways of BD SNPs were overlapped with pathways of IBD SNPs, which may be also mediated by CEBPB and PLC molecules, pivotal functional molecules in both the common pathway and regulatory SNP. These results support that intestinal BD genes shows BD phenotypes or increase intestinal BD phenotypes through complex interactions with genes with BD without intestinal involvement as well as intestinal involvement. Because the combination of multiple variants with small effects explains the overall susceptibility to multifactorial diseases and cause the same disease phenotype in complex diseases, the disruption of different biological pathways is thought to determine the intrinsic biological processes of multifactorial diseases.[27] In

this regard, these results are consistent with the studies reporting that variants in *IL10* are also associated with ulcerative colitis (UC) and Crohn's disease (CD) in infancy[63, 64] and knockout mouse of *Il10* is the best animal models of IBD, which suggest that there is a shared inherited susceptibility between BD, especially intestinal BD and IBD. In addition to implicating the association of IL10 with colon inflammation, pathway analysis showed networks beteen intestinal BD and IBD through IL10, IL12, and ERAP1. Consistently, our data showed that IL17 and IL12B were increased but IL10 reduced in inflammed colon tissues of intestinal BD and mice. It is known that key gene variants of the IL-12 and IL-23/Th17 pathway such as IL23R and STAT4. Similar to IL23R, IL12B as part of IL-23 is also a shared susceptibility gene with other IBD-associated diseases such as psoriasis[65] and ankylosing spondylitis,[66] providing an explanation of the increased incidence of these extraintestinal manifestations in IBD patients. Remarkably, the strongest cluster of common CD and UC susceptibility genes is formed by genes related to the IL-23/Th17 pathway including IL12B.[14] potential epistasis between the main IBD susceptibility genes as major pathomechanism in the disease pathogenesis. The IL12B SNP rs6887695 modulates the susceptibility and the phenotype of IBD.[67]

Consistently, *IL12B* is regareded as established IBD susceptibility gene as well as *IL23R*.[67] Biological pathway analyses in this study were in close agreement in emphasizing the crosstalk of general biological pathways

between intestinal BD and IBD or BD without intestinal involvement. Collectively, our finding strongly suggests a difference of pathway changes between intestinal BD and systemic BD, suggesting overlapping yet distinct genetic architectures for these two diseases.

Furthermore, comparisons between intestinal BD and BD without intestinal involvement showed modest evidence for differences between the two conditions in *CD180* (Lymphocyte antigen 64), *PLCB1* (phospholipase C beta 1), and *TGFBR3* (transforming growth factor, beta receptor 3), although they failed to pass Bonferroni test in the replication study (Supporting information 4). *PLCB1*, and *TGFBR3* including *DCAF12* (DDB1 and CUL4 associated factor 12), and *SCHIP1* (schwannomin interacting protein 1) were also suggested to be potential causal variants in association with intestinal BD in haplotype analysis. CD180 is a cell surface molecule consisting of extracellular leucine-rich repeats (LRR) and a short cytoplasmic tail. The extracellular LRR form the cell surface receptor complex, RP105/MD-1, which, by working in concert with toll-like receptors 4 (TLR4), controls B cell recognition and signaling of lipopolysaccharide (LPS).[68] PLCB1 is the major nuclear PLC-β isozyme and its expression is increased in the proliferating crypt compartments of the intestine in mice.[69] Interactions between the gut microbiota and epithelial cells trigger the increase of DUOX enzymatic activity via PLC-β-dependent production of inositol-1,4,5-trisphosphate.[70] *TGFBR3* had SNP of eQTL (rs1805110) and regulatory SNP

(rs17882828) in our study. Risk alleles of *TGFBR3* not only increase the expression of *IL23R*, *IL12RB2*, *KLRC4*, *TGFB1* but also decrease the expression of *IL10*. While the polymorphisms in BD were reported in other groups,[71, 72] *TGFBR3* is also known to function in colon cancer.[73, 74] *DCAF12* revealed significant increase of proinflammatory gene expression such as *CCR1*, *CCR3*, *TNF*, *IL5*, *IL13*, *ERAP1, NUDT2*[75, 76] and decrease of *IL10* expression in eQTL analyses.

V. **CONCLUSION**

Our study provides the first notable evidence that there are specific g enetic susceptibility loci associated with intestinal involvement in BD pa tients. We also found independet genetic variants associated with develo pment of intestinal BD and BD without intestinal involvement, which s howed a partial overlap with psoriasis or IBD. Of previously identified genes in GWAS of BD, *PSORS1C1* seem to be responsible for BD wit hout intestinal involvement and *ERAP1*, *ELMO1*, and *IL10* to be more responsible for intestinal BD than for BD without intestinal involvement, which suggest that there are fundamentally shared pathogenic pathways between BD and psoriasis or between BD without intestinal involvement and IBD. Our GWAS results separating between intestinal BD and BD without intestinal involvement enable a more comprehensive analysis of

disease specificity and provide insight into the common and different pa thogenic mechanisms for intestinal BD and IBD suggesting new therape utic strategies, although further larger multi-center replicated studies and additional functional experiments are required to confirm our findings.

REFERENCES

1. Ohno S, Ohguchi M, Hirose S, et al. Close association of HLA-Bw51 with Behcet's disease. Arch Ophthalmol 1982;100:1455-8.
2. Kirino Y, Bertsias G, Ishigatsubo Y, et al. Genome-wide association analysis identifies new susceptibility loci for Behcet's disease and epistasis between HLA-B*51 and ERAP1. Nat Genet 2013;45:202-7.
3. Hughes T, Coit P, Adler A, et al. Identification of multiple independent susceptibility loci in the HLA region in Behcet's disease. Nat Genet 2013;45:319-24.
4. Remmers EF, Cosan F, Kirino Y, et al. Genome-wide association study identifies variants in the MHC class I, IL10, and IL23R-IL12RB2 regions associated with Behcet's disease. Nat Genet 2010;42:698-702.
5. Mizuki N, Meguro A, Ota M, et al. Genome-wide association studies identify IL23R-IL12RB2 and IL10 as Behcet's disease susceptibility loci. Nat Genet 2010;42:703-6.
6. Hou S, Yang Z, Du L, et al. Identification of a susceptibility locus in STAT4 for Behcet's disease in Han Chinese in a genome-wide association study. Arthritis Rheum 2012;64:4104-13.
7. Cheon JH, Kim ES, Shin SJ, et al. Development and validation of novel diagnostic criteria for intestinal Behcet's disease in Korean patients with ileocolonic ulcers. Am J Gastroenterol 2009;104:2492-9.
8. Yurdakul S, Tuzuner N, Yurdakul I, et al. Gastrointestinal involvement

in Behcet's syndrome: a controlled study. Ann Rheum Dis 1996;55:208-10.

9. Gurler A, Boyvat A, Tursen U. Clinical manifestations of Behcet's disease: an analysis of 2147 patients. Yonsei Med J 1997;38:423-7.

10. Bayraktar Y, Ozaslan E, Van Thiel DH. Gastrointestinal manifestations of Behcet's disease. J Clin Gastroenterol 2000;30:144-54.

11. Chang HK, Kim JW. The clinical features of Behcet's disease in Yongdong districts: analysis of a cohort followed from 1997 to 2001. J Korean Med Sci 2002;17:784-9.

12. Bang D, Yoon KH, Chung HG, et al. Epidemiological and clinical features of Behcet's disease in Korea. Yonsei Med J 1997;38:428-36.

13. Moll JM, Haslock I, Macrae IF, et al. Associations between ankylosing spondylitis, psoriatic arthritis, Reiter's disease, the intestinal arthropathies, and Behcet's syndrome. Medicine (Baltimore) 1974;53:343-64.

14. Lees CW, Barrett JC, Parkes M, et al. New IBD genetics: common pathways with other diseases. Gut 2011;60:1739-53.

15. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. PLoS Genet 2006;2:e190.

16. Marchini J, Howie B. Genotype imputation for genome-wide association studies. Nat Rev Genet 2010;11:499-511.

17. Lake SL, Lyon H, Tantisira K, et al. Estimation and tests of haplotype-

environment interaction when linkage phase is ambiguous. Hum Hered 2003;55:56-65.

18. Dimas AS, Deutsch S, Stranger BE, et al. Common regulatory variation impacts gene expression in a cell type-dependent manner. Science 2009;325:1246-50.

19. Stranger BE, Montgomery SB, Dimas AS, et al. Patterns of cis regulatory variation in diverse human populations. PLoS Genet 2012;8:e1002639.

20. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22:1790-7.

21. Garber M, Guttman M, Clamp M, et al. Identifying novel constrained elements by exploiting biased substitution patterns. Bioinformatics 2009;25:i54-62.

22. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res 2012;40:D930-4.

23. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57-74.

24. 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;19:954-66.

25. Yang SK, 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 genetic susceptibility across ethnic populations. Gut 2014;63:80-7.

26. Xavier JM, Shahram F, Davatchi F, et al. Association study of IL10 and IL23R-IL12RB2 in Iranian patients with Behcet's disease. Arthritis Rheum 2012;64:2761-72.

27. Fridley BL, Biernacka JM. Gene set analysis of SNP data: benefits, challenges, and future directions. Eur J Hum Genet 2011;19:837-43.

28. Li Q, Winston JH, Sarna SK. Developmental origins of colon smooth muscle dysfunction in IBS-like rats. Am J Physiol Gastrointest Liver Physiol 2013;305:G503-12.

29. Shen F, Gaffen SL. Structure-function relationships in the IL-17 receptor: implications for signal transduction and therapy. Cytokine 2008;41:92-104.

30. Burgner D, Davila S, Breunis WB, et al. A genome-wide association study identifies novel and functionally related susceptibility Loci for Kawasaki disease. PLoS Genet 2009;5:e1000319.

31. Tonkin ET, Smith M, Eichhorn P, et al. A giant novel gene undergoing extensive alternative splicing is severed by a Cornelia de Lange-associated translocation breakpoint at 3q26.3. Hum Genet 2004;115:139-48.

32. 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;144:781-8.

33. Brandl K, Tomisato W, Li X, et al. Yip1 domain family, member 6 (Yipf6) mutation induces spontaneous intestinal inflammation in mice. Proc Natl Acad Sci U S A 2012;109:12650-5.

34. Happel KI, Zheng M, Young E, et al. Cutting edge: roles of Toll-like receptor 4 and IL-23 in IL-17 expression in response to Klebsiella pneumoniae infection. J Immunol 2003;170:4432-6.

35. Ruddy MJ, Wong GC, Liu XK, et al. Functional cooperation between interleukin-17 and tumor necrosis factor-alpha is mediated by CCAAT/enhancer-binding protein family members. J Biol Chem 2004;279:2559-67.

36. Fuss IJ, Becker C, Yang Z, et al. Both IL-12p70 and IL-23 are synthesized during active Crohn's disease and are down-regulated by treatment with anti-IL-12 p40 monoclonal antibody. Inflamm Bowel Dis 2006;12:9-15.

37. Su AI, Cooke MP, Ching KA, et al. Large-scale analysis of the human and mouse transcriptomes. Proc Natl Acad Sci U S A 2002;99:4465-70.

38. Fu J, Wolfs MG, Deelen P, et al. Unraveling the regulatory mechanisms underlying tissue-dependent genetic variation of gene expression. PLoS Genet 2012;8:e1002431.

39. Maldini C, Lavalley MP, Cheminant M, et al. Relationships of HLA-B51 or B5 genotype with Behcet's disease clinical characteristics: systematic review and meta-analyses of observational studies. Rheumatology (Oxford) 2012;51:887-900.

40. Van Limbergen J, Wilson DC, Satsangi J. The genetics of Crohn's disease. Annu Rev Genomics Hum Genet 2009;10:89-116.

41. de Menthon M, Lavalley MP, Maldini C, et al. HLA-B51/B5 and the risk of Behcet's disease: a systematic review and meta-analysis of case-control genetic association studies. Arthritis Rheum 2009;61:1287-96.

42. Yasuoka H, Yamaguchi Y, Mizuki N, et al. Preferential activation of circulating CD8+ and gammadelta T cells in patients with active Behcet's disease and HLA-B51. Clin Exp Rheumatol 2008;26:S59-63.

43. Ombrello MJ, Kirino Y, de Bakker PI, et al. Behcet disease-associated MHC class I residues implicate antigen binding and regulation of cell-mediated cytotoxicity. Proc Natl Acad Sci U S A 2014;111:8867-72.

44. Sham PC, Purcell SM. Statistical power and significance testing in large-scale genetic studies. Nat Rev Genet 2014;15:335-46.

45. Doran G. The short and the long of UTRs. J RNAi Gene Silencing 2008;4:264-5.

46. Halvorsen M, Martin JS, Broadaway S, et al. Disease-associated mutations that alter the RNA structural ensemble. PLoS Genet 2010;6:e1001074.

47. Wallace GR, Kondeatis E, Vaughan RW, et al. IL-10 genotype analysis in patients with Behcet's disease. Hum Immunol 2007;68:122-7.

48. Genetic Analysis of Psoriasis C, the Wellcome Trust Case Control C, Strange A, et al. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet 2010;42:985-90.

49. van Ham TJ, Kokel D, Peterson RT. Apoptotic cells are cleared by directional migration and elmo1- dependent macrophage engulfment. Curr Biol 2012;22:830-6.

50. Haroon N, Inman RD. Endoplasmic reticulum aminopeptidases: Biology and pathogenic potential. Nat Rev Rheumatol 2010;6:461-7.

51. Saveanu L, Carroll O, Lindo V, et al. Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. Nat Immunol 2005;6:689-97.

52. Yan J, Parekh VV, Mendez-Fernandez Y, et al. In vivo role of ER-associated peptidase activity in tailoring peptides for presentation by MHC class Ia and class Ib molecules. J Exp Med 2006;203:647-59.

53. Azuma YT, Matsuo Y, Kuwamura M, et al. Interleukin-19 protects mice from innate-mediated colonic inflammation. Inflamm Bowel Dis 2010;16:1017-28.

54. Wang MH, Fiocchi C, Ripke S, et al. A novel approach to detect cumulative genetic effects and genetic interactions in Crohn's disease.

Inflamm Bowel Dis 2013;19:1799-808.

55. Yamamoto-Furusho JK, Alvarez-Leon E, Fragoso JM, et al. Protective role of interleukin-19 gene polymorphisms in patients with ulcerative colitis. Hum Immunol 2011;72:1029-32.

56. Fonseca-Camarillo G, Furuzawa-Carballeda J, Granados J, et al. Expression of interleukin (IL)-19 and IL-24 in inflammatory bowel disease patients: a cross-sectional study. Clin Exp Immunol 2014;177:64-75.

57. Robinson PC, Claushuis TA, Cortes A, et al. Genetic dissection of acute anterior uveitis reveals similarities and differences in associations observed with ankylosing spondylitis. Arthritis Rheumatol 2014.

58. Berkhout M, Friederich P, van Krieken JH, et al. Low detoxification capacity in the ileal pouch mucosa of patients with ulcerative colitis. Inflamm Bowel Dis 2006;12:112-6.

59. Olszak T, Neves JF, Dowds CM, et al. Protective mucosal immunity mediated by epithelial CD1d and IL-10. Nature 2014;509:497-502.

60. Mittal RD, Manchanda PK, Bid HK, et al. Analysis of polymorphisms of tumor necrosis factor-alpha and polymorphic xenobiotic metabolizing enzymes in inflammatory bowel disease: study from northern India. J Gastroenterol Hepatol 2007;22:920-4.

61. Elder JT. PSORS1: linking genetics and immunology. J Invest Dermatol 2006;126:1205-6.

62. Gambichler T, Kreuter A, Susok L, et al. Glutathione-S-transferase T1 genotyping and phenotyping in psoriasis patients receiving treatment with oral fumaric acid esters. J Eur Acad Dermatol Venereol 2014;28:574-80.

63. Franke A, Balschun T, Karlsen TH, et al. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. Nat Genet 2008;40:1319-23.

64. Franke A, McGovern DP, Barrett JC, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 2010;42:1118-25.

65. Cargill M, Schrodi SJ, Chang M, et al. A large-scale genetic association study confirms IL12B and leads to the identification of IL23R as psoriasis-risk genes. Am J Hum Genet 2007;80:273-90.

66. Wellcome Trust Case Control C, Australo-Anglo-American Spondylitis C, Burton PR, et al. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet 2007;39:1329-37.

67. Glas J, Seiderer J, Wagner J, et al. Analysis of IL12B gene variants in inflammatory bowel disease. PLoS One 2012;7:e34349.

68. Miura Y, Shimazu R, Miyake K, et al. RP105 is associated with MD-1 and transmits an activation signal in human B cells. Blood 1998;92:2815-22.

69. Lee SJ, Leoni G, Neumann PA, et al. Distinct phospholipase C-beta isozymes mediate lysophosphatidic acid receptor 1 effects on intestinal epithelial homeostasis and wound closure. Mol Cell Biol 2013;33:2016-28.

70. Ha EM, Lee KA, Park SH, et al. Regulation of DUOX by the Galphaq-phospholipase Cbeta-Ca2+ pathway in Drosophila gut immunity. Dev Cell 2009;16:386-97.

71. Chen Y, Yang P, Li F, et al. Association analysis of TGFBR3 gene with Vogt-Koyanagi-Harada disease and Behcet's disease in the Chinese Han population. Curr Eye Res 2012;37:312-7.

72. Hou S, Kijlstra A, Yang P. The genetics of Behcet's disease in a Chinese population. Front Med 2012;6:354-9.

73. Yan Z, Deng X, Friedman E. Oncogenic Ki-ras confers a more aggressive colon cancer phenotype through modification of transforming growth factor-beta receptor III. J Biol Chem 2001;276:1555-63.

74. Gatza CE, Holtzhausen A, Kirkbride KC, et al. Type III TGF-beta receptor enhances colon cancer cell migration and anchorage-independent growth. Neoplasia 2011;13:758-70.

75. Hankin S, Matthew N, Thorne H, et al. Diadenosine 5',5'''-P1,P4-tetraphosphate hydrolase is present in human erythrocytes, leukocytes and platelets. Int J Biochem Cell Biol 1995;27:201-6.

76. Carmi-Levy I, Yannay-Cohen N, Kay G, et al. Diadenosine tetraphosphate hydrolase is part of the transcriptional regulation network in immunologically activated mast cells. Mol Cell Biol 2008;28:5777-84.

<ABSTRACT (IN KOREAN)>

유전체 전장 연관성 분석에 의한 베체트 장염의 감수성 유전자 발굴

<지도교수 천 재 희>

연세대학교 대학원 의학과

정 윤 숙

최근 몇몇 유전체 전장 연관성 분석 연구들에 의해서 베체트병과 염증성 장질환의 발생과 관련이 있는 유전자들이 밝혀지고 있다. 하지만 현재까지 베체트 장염의 발생과 관련 있는 유전자들에 대한 연구는 시행된 바가 없다. 본 연구에서는 유전체 전장 연관성 분석을 통하여 베체트 장염의 감수성 유전자를 조사하였다.

100명의 장질환 침범이 없는 베체트병군과 99명의 베체트 장염군, 그리고 557명의 정상 대조군을 대상으로 유전체 전장 연관성 분석을 시행하였다. 새로운 코호트인 138명의 장질환 침범이 없는 베체트병군과 196명의 베체트

장염군을 대상으로 후보 유전자들 변이를 검증하였다 (validation study). 또한 후보 유전자들의 단백질 발현을 평가하기 위해 면역조직화학 염색을 시행하였다.

유전체 전장 연관성 분석과 검증 연구 (validation study) 를 통해 *NAALADL2* (rs3914501, OR = 1.914, P = 3.5 × 10−4 in a recessive model) 유전자가 베체트 장염의 감수성 유전자임을 확인할 수 있었다. 또한 *NAALADL2* 유전자는 베체트 장염의 발생뿐만 아니라 누적 수술률, 입원률, 스테로이드 사용률과도 관련이 있어 나쁜 예후와도 연관성이 있었다. *ERAP1* 유전자와 *IL10* 유전자 역시 베체트 장염의 발생과 연관성이 있었다. 한편, 장질환 침범이 없는 베체트병군과 베체트 장염군간의 비교 분석 결과, *YIPF7* (rs6838327, OR = 1.522, P = 5.6 × 10−4) 유전자 변이가 의미있게 차이가 있었다. 장질환 침범이 없는 베체트병의 감수성 유전자는 이미 기존 연구에서 베체트병과 관련성이 보고된 바 있는 *PSORS1C1* (rs12525170, OR = 2.210, P = 1.1 × 10−5) 유전자임 확인할 수 있었다.

결론적으로 본 연구를 통해 세계 최초로 베체트 장염의

감수성 유전자들을 발굴할 수 있었다. 이런 발굴된 유전자들 일부는 염증성 장질환이나 다른 자가 면역 질환들과도 연관성이 보고된 바 있다. 본 연구 결과들은 베체트 장염이 다른 자가 면역 질환들과 구별되는 유전적 특징을 가지고 있을 뿐 아니라 염증성 징질환이나 다른 자가 면역질환과 유전적으로 공유되는 특징도 가지고 있음을 시사한다.

핵심되는 말 : 베체트 장염, 유전체 전장 연관성 분석 연구, 질병 감수성 유전자