

# Strong Evidence of a Combination Polymorphism of the Tyrosine Kinase 2 Gene and the Signal Transducer and Activator of Transcription 3 Gene as a DNA-Based Biomarker for Susceptibility to Crohn's Disease in the Japanese Population

Kayoko Sato · Mizuho Shiota · Sayaka Fukuda · Eiko Iwamoto · Haruhisa Machida · Tatsuo Inamine · Shinji Kondo · Katsunori Yanagihara · Hajime Isomoto · Yohei Mizuta · Shigeru Kohno · Kazuhiro Tsukamoto

Received: 28 June 2009 / Accepted: 14 July 2009 / Published online: 4 August 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

## Abstract

**Objective** An association between susceptibility to inflammatory bowel disease (IBD) and polymorphisms of both the tyrosine kinase 2 gene (*TYK2*) and the signal transducer and activator of transcription 3 gene (*STAT3*) was examined in a Japanese population in order to identify the genetic determinants of IBD.

**Methods** The study subjects comprised 112 patients with ulcerative colitis, 83 patients with Crohn's disease (CD), and 200 healthy control subjects. Seven tag single-nucleotide polymorphisms (SNPs) in *TYK2* and *STAT3* were detected by PCR-restriction fragment length polymorphism.

**Results** The frequencies of a C allele and its homozygous C/C genotype at rs2293152 SNP in *STAT3* in CD patients were significantly higher than those in control subjects

( $P=0.007$  and  $P=0.001$ , respectively). Furthermore, out of four haplotypes composed of the two tag SNPs (rs280519 and rs2304256) in *TYK2*, the frequencies of a Hap 1 haplotype and its homozygous Hap 1/Hap1 diplotype were significantly higher in CD patients in comparison to those in control subjects ( $P=0.023$  and  $P=0.024$ , respectively). In addition, the presence of both the C/C genotype at rs2293152 SNP in *STAT3* and the Hap 1/Hap 1 diplotype of *TYK2* independently contributes to the pathogenesis of CD and significantly increases the odds ratio to 7.486 for CD ( $P=0.0008$ ).

**Conclusion** *TYK2* and *STAT3* are genetic determinants of CD in the Japanese population. This combination polymorphism may be useful as a new genetic biomarker for the identification of high-risk individuals susceptible to CD.

K. Sato · M. Shiota · S. Fukuda · E. Iwamoto · T. Inamine · S. Kondo · K. Tsukamoto (✉)  
Department of Pharmacotherapeutics, Nagasaki University  
Graduate School of Biomedical Sciences,  
1-14 Bunkyo-machi,  
Nagasaki 852-8521, Japan  
e-mail: ktsuka@nagasaki-u.ac.jp

H. Machida · K. Yanagihara · H. Isomoto · Y. Mizuta · S. Kohno  
Second Department of Internal Medicine,  
Nagasaki University School of Medicine,  
Sakamoto 1-7-1,  
Nagasaki 852-8501, Japan

K. Tsukamoto  
AVSS, Co., Ltd,  
1-22 Wakaba-machi,  
Nagasaki 852-8137, Japan

**Keywords** *TYK2* · *STAT3* · polymorphisms · Crohn's disease · candidate gene-based association study · DNA-based biomarker · Japanese population

## Introduction

Idiopathic inflammatory bowel disease (IBD) is a multifactorial disorder characterized by chronic and relapsing inflammation specific to the gastrointestinal tract, thus resulting in intestinal malabsorption, mucosal immune system abnormalities, and exaggerated inflammatory responses [1–5]. IBD has two main subtypes, namely ulcerative colitis (UC) and Crohn's disease (CD). Although the precise etiology remains unknown, several environmen-

tal factors, such as commensal bacteria, food antigens, and smoking, as well as multiple genetic factors may contribute to the occurrence and development of IBD [1–5]. Genome-wide linkage-based family studies, candidate gene-based association studies, and large-scale genome-wide association (GWA) studies using single-nucleotide polymorphisms (SNPs) have shown possible IBD susceptibility genes and chromosomal loci [6–10].

IBD is involved in a complex interplay of innate and adaptive immune cells, including lymphocytes, macrophages, and dendritic cells. In such a setting, alterations in cytokine synthesis and cytokine signaling pathways are attributed to the pathogenesis of IBD [1–5]. GWA studies have recently indicated that the multiple genes implicated in the interleukin 23 (IL-23) and its receptor (IL23R) signaling pathway are associated with susceptibility to CD as well as UC [6–12]. For example, IL-23, IL23R, interleukin 12 precursor (IL12B), interleukin 12 receptor (IL12R), the Janus kinase (JAK) families, and the signal transducers and activators of transcription (STAT) families belong to a gene network in the IL-23/IL23R signaling pathway [13], implying that a subset of these genes can play a central role in the pathogenesis of IBD and may function as a key conductor of innate and adaptive inflammatory responses at multiple levels in the intestinal mucosa of IBD patients.

*IL-23* as well as *JAK2* [9, 14] and *STAT3* [9, 12, 14] are associated with susceptibility to CD. These genes are involved in a gene network in the IL23/IL23R signaling pathway. *JAK2* has been recently identified as a CD susceptibility gene in a meta-analysis of GWA data [9, 14] but not by any single-marker association studies on CD. Likewise, tyrosine kinase 2 (*TYK2*), which is a member of the JAK families located in  $\beta 1$  subunit of IL12 receptor (IL12RB1) [4, 15–17] and in gp130 of interleukin 6 receptor (IL6R) [18], is also identified as a CD susceptibility gene by the meta-analysis [10] but did not reach significance by single-marker association studies. Furthermore, *TYK2* is activated via signaling from a broader range of cytokine receptors and induces phosphorylation, homodimerization, and nuclear translocation of *STAT3* [15–19], thus resulting in several gene transcriptions and leading to IL-23-induced production of IL-17, a pro-inflammatory cytokine in natural killer cells, natural killer T cells, CD4<sup>+</sup> T cells, and CD8<sup>+</sup> T cells [20–22]. This signaling cascade also plays a role in the differentiation of CD4<sup>+</sup> (naive) T cells into Th17 cells [17, 20–22], which is involved in the first line of host defense by controlling immune responses [22].

Therefore, we performed a candidate gene-based association study by selecting *TYK2* and *STAT3* as candidate genes. The purpose of this study was to investigate whether SNPs and their combination polymorphisms, which are referred to as haplotypes, in *TYK2*

and *STAT3* are also associated with susceptibility to IBD in a Japanese population and whether such polymorphisms can be used as new genetic biomarkers for predicting the onset of IBD.

## Subjects and Methods

### Subjects

The study subjects were Japanese who were unrelated to one another. The subjects included 112 patients with UC, 83 patients with CD, and 200 gender-matched, healthy volunteers as control subjects. IBD patients were enrolled from eight general hospitals in Nagasaki, Japan from October 2003 to October 2008. The clinical characteristics of the subjects at the end point of this study are shown in Table I. The study protocol was approved by the Committee for Ethical Issue dealing with the Human Genome and Gene Analysis at Nagasaki University, and written informed consent was obtained from each subject.

The diagnosis of IBD was made based on the endoscopic, radiological, histological, and clinical criteria established by both the World Health Organization Council for International Organizations of Medical Sciences and the International Organization for the Study of Inflammatory Bowel Disease [23–25]. Patients with indeterminate colitis, multiple sclerosis, systemic lupus erythematosus, or any other diagnosed autoimmune diseases were excluded from this study.

Patients with UC were classified into subgroups according to age at onset ( $\leq 40$  or  $\geq 40$  years), extension of disease (proctitis, left-sided colitis, or pancolitis), disease severity (mild, moderate, or severe), and disease activity (active or inactive) (Table I). Likewise, patients with CD were classified into subgroups according to age at onset ( $\leq 40$  or  $\geq 40$  years), the location of lesions (ileal, ileocolonic, colonic, or isolated upper), disease severity (mild, moderate, or severe), disease activity (active or inactive), and the behavior of disease (stricturing, penetrating, or perianal; Table I). The location and extension of UC and CD, disease severity of UC and CD, and behavior of CD were stratified in accordance with the Montreal classification [26] with slight modification. A high clinical activity index (CAI $\geq 5$ ) for UC [27] and a high Crohn's disease activity index (CDAI $\geq 150$ ) [28] were regarded as active-phase patients.

### Preparation of Genomic DNA

Genomic DNA was extracted from a whole blood sample from each subject using a DNA Extractor WB-Rapid Kit

**Table 1** The Clinical Characteristics of Study Subjects

Characteristics	Patients with		Control subjects
	UC	CD	
Number	114	83	200
Age, mean±SD (years)	44.2±16.7*	34.3±12.5	32.5±11.2
Age range (years)	14–83	17–75	20–60
Male/female (%)	59/55 (51.7/48.3)	50/33 (60.2/39.8)	126/74 (62.5/37.5)
Age at onset			
≤40 years	39	11	
≥40 years	75	72	
Extent of UC			
Proctitis	14		
Left-sided colitis	43		
Pancolitis	57		
Location of CD			
Ileal		16	
Ileocolonic		55	
Colonic		11	
Isolated upper		1	
Disease severity			
Mild	51	17	
Moderate	38	45	
Severe	21	9	
Unknown	4	12	
Disease activity			
Active	63	54	
Inactive	48	17	
Unknown	3	12	
Behavior of CD <sup>a</sup>			
Stricturing		44	
Penetrating		40	
Perianal diseases		36	

SD standard deviation

\* $P < 0.01$  in comparison to control subjects

<sup>a</sup>Number of the affected patients with CD

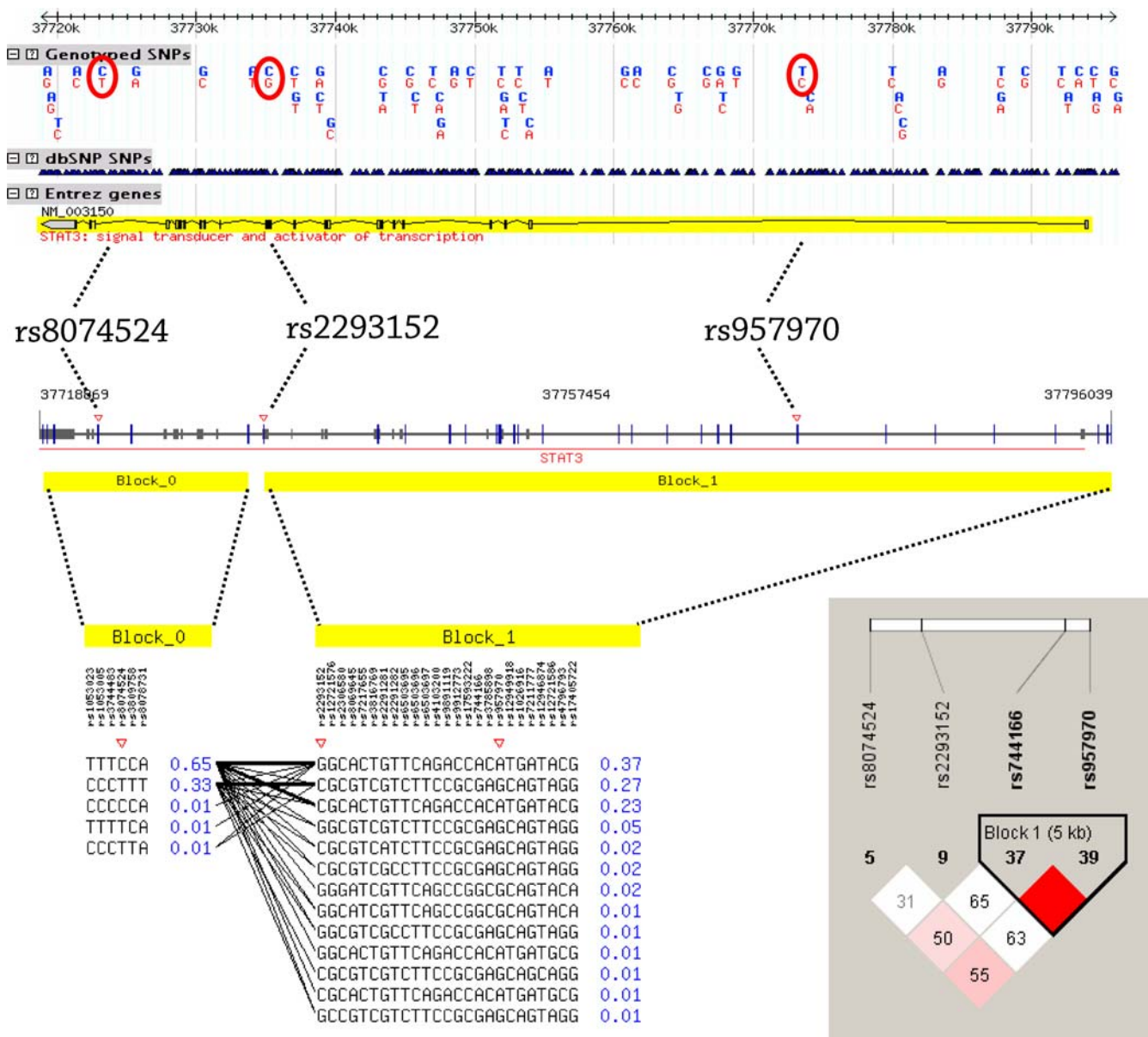
(Wako, Osaka, Japan) according to the manufacturer's protocol.

#### Sources of the Candidate Genes and Their Polymorphisms

All of SNPs in *STAT3* (GenBank accession number, AY572796; MIM 102582) located on chromosome 17q21 [29] and *TYK2* (GenBank accession number, AY549314; MIM 176941) located on chromosome 19p13.2 [30] were obtained using data available on the International HapMap Web site [31]. Candidate tag SNPs were selected with priority in a minor allele frequency of more than 5%. Subsequently, linkage disequilibrium blocks and genotyped tag SNPs among the candidate tag SNPs were determined using the iHAP software program [32]. The gene structure and positions of the genotyped tag SNP sites in *STAT3* and *TYK2* are shown in Figs. 1 and 2, respectively.

#### Determination of Three SNPs in *STAT3*

Three SNPs, rs8074524 in intron 3, rs2293152 in intron 11, and rs957970 in intron 23, were selected as genotyped tag SNPs (Fig. 1) and were subsequently analyzed by polymerase chain reaction (PCR)-restriction fragment length polymorphism (RFLP). The polymorphic region was amplified by PCR with a GeneAmp PCR System 9700 thermal cycler (Applied Biosystems, Foster City, CA, USA) using 25 ng of genomic DNA in a 25- $\mu$ l reaction mixture containing 0.8 $\times$  GoTaq Green master mix (Promega, Madison, WI, USA) and 15 pmol each of the following primers: forward primer 5'-GTCTGGAAAGCTC CATCTGC-3' and reverse primer 5'-AGAGGCCAGAT TAGTGCTGG-3' for rs8074524; forward primer 5'-TCCCCTGTGATTGATCCCC-3' and reverse primer 5'-CATTCCCACATCTCTGCTCC-3' for rs2293152; and forward primer 5'-CTGGGCTCAAGTGATCTTCC-3' and



**Fig. 1** Locations of the genotyped tag SNP sites in *STAT3* in the International HapMap (upper) and iHap (lower) data. The horizontal bars in the middle indicate the genomic sequence of *STAT3*. Blue vertical bars indicate the positions of all SNP sites. Yellow rectangles represent the positions of linkage disequilibrium blocks. A list and the

locations of candidate tag SNPs are shown in the lower right position using Haploview 4.0 software. Red inverted triangles indicate the genotyped tag SNPs sites in this study, and their names are presented above each inverted triangle

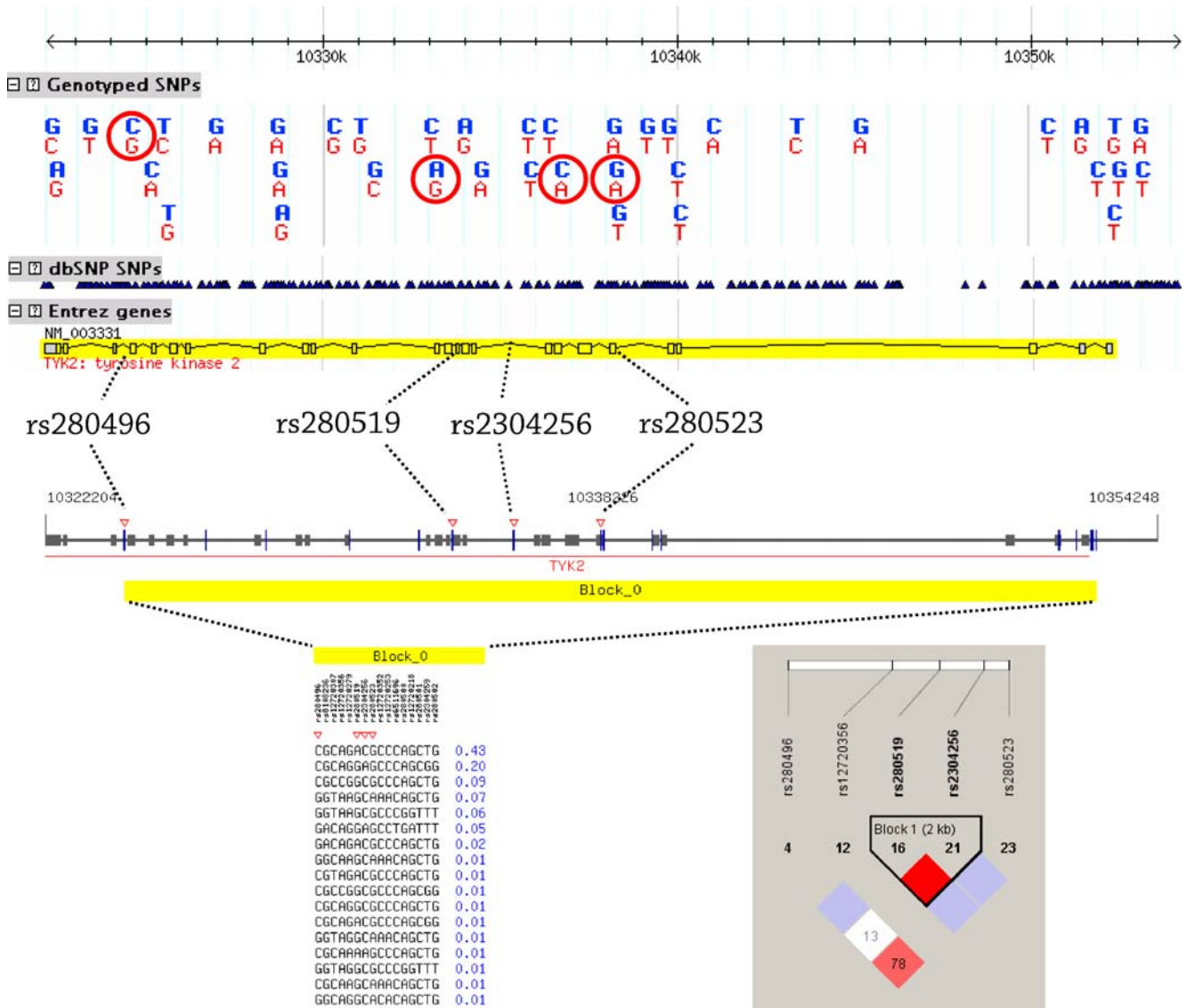
reverse primer 5'-GTACTCATCGCCCTCCATTG-3' for rs957970. The amplification protocol comprised initial denaturation at 95°C for 2 min, followed by 30 cycles of denaturation at 95°C for 30 s, annealing at 62°C for 30 s, and extension at 72°C for 30 s and final extension at 72°C for 5 min. The PCR products were digested with restriction enzyme, *Hpa* II (Takara Bio Inc., Kyoto, Japan) for rs8074524 and rs2293152; *Xsp* I (Takara Bio Inc.) for rs957970. The digests were separated by electrophoresis on a 2% agarose gel (Nacalai Tesque, Kyoto, Japan) and

visualized with an ultraviolet transilluminator (Alpha Innotech Co., San Leandro, CA, USA) after ethidium bromide (Nacalai Tesque) staining.

#### Determination of Four SNPs in *TYK2*

Four genotyped tag SNPs in *TYK2*, rs280496 in intron 3, rs280519 in intron 14, rs2304256 in intron 18, and rs280523 in intron 20 (Fig. 2), were analyzed by PCR-RFLP using 15 pmol each of the following primers:





**Fig. 2** Locations of the genotyped tag SNP sites in *TYK2* in the International HapMap (upper) and iHap (lower) data. The horizontal bars in the middle indicate the genomic sequence of *TYK2*. Blue vertical bars indicate the positions of all SNP sites. A yellow rectangle represents the positions of a linkage disequilibrium block. A list and

the locations of candidate tag SNPs are shown in the lower right position using Haploview 4.0 software. Red inverted triangles indicate the genotyped tag SNPs sites in this study, and their names are presented above each inverted triangle

forward primer 5'-CGGGGTGATATGCTCATTGG-3' and reverse primer 5'-CAACGTGCTGCTGGACAACG-3' for rs280496; forward primer 5'-CCGCCATGGTGAAAGT TAGC-3' and reverse primer 5'-ATTTGTGCAGGC CAAGCTGC-3' for rs280519; forward primer 5'-TCAC CAGGCACTTGTGTGCC-3' and reverse primer 5'-CGGCTTCCAGCATGTGTATG-3' for rs2304256; and forward primer 5'-ACATTTCCCCCTGCCTACAC-3' and reverse primer 5'-TTACAGACATGCGCCACCAC-3' for rs280523. The other constituents of the PCR mixture were the same as described above. The amplification protocol

comprised initial denaturation at 95°C for 2 min, followed by 30 cycles of denaturation at 95°C for 30 s, annealing at 64°C (rs280496, rs280519, and rs280523) or 62°C (rs2304256) for 30 s, and extension at 72°C for 30 s and final extension at 72°C for 5 min. The PCR products were digested with *Bsl* I (New England BioLabs Inc., Beverly, MA, USA) for rs280496, *Hpy*99 I (New England BioLabs Inc.) for rs280519, *Bsm* I (New England BioLabs Inc.) for rs2304256, and *BsiE* I (New England BioLabs Inc.) for rs280523. The digests were then separated on a 2% agarose gel as described above.

### Haplotype Structure of *TYK2*

Two tag SNPs in *TYK2*, which showed a close association of susceptibility to CD and were located within the same linkage disequilibrium block (Fig. 2), were utilized to infer the haplotype structure as well as to analyze the haplotype frequency using the SNP Alyze 7.0 standard software package (Dynacom Inc., Yokohama, Japan) to emphasize the variability and to enhance the power of detecting allelic association of rare variants [33, 34].

### Statistical Analysis

Differences in age and gender between UC or CD patients and control subjects were evaluated by an unpaired Student's *t* test and a chi-square test, respectively, using the SPSS 17 (SPSS Japan Inc., Tokyo, Japan) and Prism 5 (GraphPad Software Inc., San Diego, CA, USA) statistical software packages. The frequencies of the expected alleles were calculated from those of the observed genotypes according to the Hardy–Weinberg equilibrium. The frequencies of the observed and expected alleles were compared by the chi-square test with Yates' correction using the SNP Alyze 7.0 standard software package. The frequencies and distributions of alleles, genotypes, haplotypes, and diplotypes were statistically compared between UC or CD patients and control subjects by the chi-square test and logistic regression analysis using Prism 5 and SPSS 17. Subsequently, a comparison of the genetic risk factors between the statistically significant genotype of *STAT3* and

diplotype of *TYK2* for susceptibility to CD was carried out by a multivariate logistic regression analysis using SPSS 17. The odds ratio (OR) with 95% confidence interval (CI) was calculated using SPSS 17. A *P* value of less than 0.05 was considered to be statistically significant.

### Results

#### Association of Tag SNPs in *STAT3* with Susceptibility to IBD

The frequencies and distributions of alleles and genotypes at the three tag SNPs in *STAT3* were identified and compared between UC or CD patients and control subjects (Tables II and III, respectively). The C allele at rs8074524 SNP, G allele at rs2293152 SNP, and A allele at rs957970 SNP are major alleles, whereas the other alleles are minor alleles (Table II). The distributions of these tag SNPs in *STAT3* among IBD patients and control subjects corresponded well to the Hardy–Weinberg equilibrium, thus implying that the subject base has a homogeneous genetic background.

The frequencies of the C allele and its homozygous C/C genotype at rs2293152 in CD patients were significantly higher than those in control subjects (45.2% vs. 33.3%, *P*=0.007 and 22.9% vs. 8.5%, *P*=0.001, respectively). No significant differences were observed in the frequency of other alleles and genotypes between patients and control subjects.

**Table II** Distributions of Polymorphic Alleles at the Genotyped Tag SNP Sites in *STAT3* and *TYK2* Among Study Subjects

Gene	SNP	Allele	Number (%) of alleles in		Allele comparison <sup>a</sup> <i>P</i> value	Number (%) of alleles in		Allele comparison <sup>a</sup> <i>P</i> value
			UC	Control		CD	Control	
<i>STAT3</i>	rs8074524	C	144 (64.3)	253 (63.3)	0.759	106 (63.9)	253 (63.3)	0.892
		T	80 (35.7)	147 (36.7)		60 (36.1)	147 (36.7)	
	rs2293152	G	151 (66.2)	267 (66.7)	0.894	91 (54.8)	267 (66.7)	0.007
		C	77 (33.8)	133 (33.3)		75 (45.2)	133 (33.3)	
	rs957970	A	116 (51.8)	218 (54.5)	0.652	84 (50.6)	218 (54.5)	0.397
		G	108 (48.2)	182 (45.5)		82 (49.4)	182 (45.5)	
<i>TYK2</i>	rs280496	C	201 (89.7)	362 (90.5)	0.784	144 (86.7)	362 (90.5)	0.187
		G	23 (10.3)	38 (9.5)		22 (13.3)	38 (9.5)	
	rs280519	A	136 (60.7)	224 (56.0)	0.261	109 (65.7)	224 (56.0)	0.034
		G	88 (39.3)	176 (44.0)		57 (34.3)	176 (44.0)	
	rs2304256	C	156 (69.6)	262 (65.5)	0.311	129 (77.7)	262 (65.5)	0.004
		A	68 (30.4)	138 (34.5)		37 (23.3)	138 (34.5)	
	rs280523	G	207 (92.4)	370 (92.5)	0.968	154 (92.8)	370 (92.5)	0.911
		A	17 (7.6)	30 (7.5)		12 (7.2)	30 (7.5)	
Total number of alleles			228	400		166	400	

<sup>a</sup> Each allele was compared to another allele using a chi-square test.

**Table III** Distribution of Genotypes at the Tag SNP Sites in *STAT3* and *TYK2* Between CD Patients and Control Subjects

Gene	SNP	Genotype	Number (%) of genotypes in		Genotype comparison <sup>a</sup>	
			CD	Control	OR (95% CI)	P value
<i>STAT3</i>	rs8074524	C/C	34 (41.0)	75 (37.5)	1.156 (0.686–1.951)	0.586
		C/T	38 (45.8)	103 (51.5)	0.795 (0.476–1.329)	0.382
		T/T	11 (13.3)	22 (11.0)	1.236 (0.570–2.680)	0.591
	rs2293152	G/G	27 (32.5)	84 (42.0)	0.666 (0.389–1.141)	0.139
		G/C	37 (44.6)	99 (49.5)	0.821 (0.491–1.372)	0.451
		C/C	19 (22.9)	17 (8.5)	3.196 (1.566–6.523)	0.001
	rs957970	A/A	19 (22.9)	55 (27.5)	0.783 (0.430–1.424)	0.423
		G/A	46 (55.4)	108 (54.0)	1.059 (0.633–1.772)	0.827
		G/G	18 (21.7)	37 (18.5)	1.220 (0.648–2.296)	0.538
<i>TYK2</i>	rs280496	C/C	61 (73.5)	162 (81.0)	0.650 (0.356–1.187)	0.161
		C/G	22 (26.5)	38 (19.0)	1.538 (0.842–2.807)	0.161
		G/G	0	0	–	–
	rs280519	A/A	37 (44.6)	63 (31.5)	1.749 (1.034–2.959)	0.037
		A/G	35 (42.2)	98 (49.0)	0.759 (0.453–1.272)	0.295
		G/G	11 (13.2)	39 (19.5)	0.631 (0.306–1.302)	0.212
	rs2304256	C/C	52 (62.7)	86 (43.0)	2.224 (1.315–3.716)	0.003
		C/A	25 (30.1)	90 (45.0)	0.527 (0.305–0.909)	0.021
		A/A	6 (7.2)	24 (12.0)	0.571 (0.225–1.454)	0.240
	rs280523	G/G	71 (85.5)	170 (85.0)	1.044 (0.506–2.155)	0.907
		G/A	12 (14.5)	30 (15.0)	0.958 (0.464–1.976)	0.907
		A/A	0	0	–	–
Total number of subjects			83	200		

OR odds ratio, CI confidence interval

<sup>a</sup> Each genotype was compared to other genotypes combined using a logistic regression analysis

Association of Tag SNPs in *TYK2* with Susceptibility to IBD

The frequencies and distributions of alleles and genotypes at the four tag SNPs in *TYK2* were identified and compared between UC or CD patients and control subjects (Tables II and III, respectively). The C allele at rs280496 SNP, A allele at rs280519 SNP, C allele at rs2304256 SNP, and G allele at rs280523 SNP are major alleles, whereas other alleles are minor alleles (Table II). The distributions of these tag SNPs in *TYK2* among IBD patients and control subjects corresponded well to the Hardy–Weinberg equilibrium. The frequencies of the A allele and its homozygous A/A genotype at rs280519 in CD patients were significantly higher than those in control subjects (65.7% vs. 56.0%,  $P=0.034$  and 44.6% vs. 31.5%,  $P=0.037$ , respectively). Likewise, the frequencies of the C allele and its homozygous C/C genotype at rs2304256 in CD patients were also significantly higher than those in control subjects (77.7% vs. 65.5%,  $P=0.004$  and 62.7% vs. 43.0%,  $P=0.003$ , respectively). In contrast, the frequency of the C/A heterozygous genotype at rs2304256 was significantly lower in CD patients in comparison to that in control subjects (30.1% vs. 45.0%,  $P=0.021$ ).

Association of Haplotypes and Diplotypes of *TYK2* with Susceptibility to CD

Subsequently, four haplotypes composed of these two tag SNPs (rs280519 and rs2304256), which displayed a significant association with CD susceptibility and were located within the same linkage disequilibrium block, were constructed and identified using the SNP Alyze 7.0 standard software package (Table IV). A logistic regression analysis revealed the frequency of a Hap 1 haplotype (A allele at rs280519 SNP and C allele at rs2304256 SNP) to significantly increase in CD patients in comparison to that in control subjects (65.7% vs. 55.3%,  $P=0.023$ , OR=1.549). In contrast, the frequency of a Hap 2 haplotype (G allele at rs280519 SNP and A at rs2304256 SNP) was significantly decreased in CD patients in comparison to that in control subjects (22.3% vs. 33.7%,  $P=0.007$ , OR=0.563).

Furthermore, eight diplotypes composed of four haplotypes were identified (Table V). A logistic regression analysis showed that the frequency of the CD patients possessing a Hap 1/Hap 1 diplotype was significantly higher than that of the control subjects (44.6% vs. 30.5%,  $P=0.024$ , OR=1.833). In contrast, the frequency of the CD

**Table IV** Distributions of Haplotypes of *TYK2* Between CD Patients and Control Subjects

Haplotype	SNP		Number (%) of haplotypes in		Haplotype comparison <sup>a</sup>	
	rs280519	rs2304256	CD	Control	OR (95% CI)	<i>P</i> value
Hap 1	A	C	109 (65.7)	221 (55.3)	1.549 (1.063–2.256)	0.023
Hap 2	G	A	37 (22.3)	135 (33.7)	0.563 (0.370–0.857)	0.007
Hap 3	G	C	20 (12.0)	41 (10.3)	1.199 (0.680–2.117)	0.530
Hap 4	A	A	0	3 (0.7)	–	–
Total number of haplotypes			166	400		

OR odds ratio, CI confidence interval

<sup>a</sup> Each haplotype was compared to other haplotypes combined using a logistic regression analysis

patients having a Hap 1/Hap 2 diplotype was significantly lower than that of the control subjects (24.1% vs. 36.5%,  $P=0.045$ , OR=0.552). The results of the diplotype analysis regarding the Hap 1 haplotype of *TYK2* coincided well with those of the haplotype analysis between CD patients and control subjects; however, a Hap 2/Hap 2 diplotype showed no statistically significant lack of susceptibility to CD (Table V).

#### Gene–Gene Interaction Between *STAT3* and *TYK2* for Susceptibility to CD

The gene–gene interaction between *STAT3* and *TYK2* was analyzed between CD patients and control subjects. A multivariate logistic regression analysis indicated that two variable genetic factors, the C/C genotype at rs2293152 SNP in *STAT3* and the Hap 1/Hap 1 diplotype of *TYK2*, independently contributed to susceptibility to CD ( $P=0.002$ , OR=3.113, 95% CI=1.515–6.399 and  $P=0.030$ , OR=1.783, 95% CI=1.042–3.053, respectively; Table VI).

Furthermore, with regard to the gene–gene combination effect of *STAT3* genotype and *TYK2* diplotype for susceptibility to CD, a multivariate logistic regression analysis showed the OR to significantly increase (7.486,  $P=0.0008$ , 95% CI=2.310–24.261) in the individuals possessing both the C/C genotype at rs2293152 SNP in *STAT3* and the Hap

1/Hap 1 diplotype of *TYK2* in comparison to that in the individuals possessing the other genotypes (Table VII).

#### Discussion

This study is the first demonstration of the single-marker association of *STAT3* and *TYK2* polymorphisms with CD susceptibility in the Japanese population, although meta-analyses using GWA data previously indicated that *STAT3* and *TYK2* appear to be the genetic determinants of CD in the European and North American populations [9, 10, 14]. Furthermore, *STAT3* is associated with only CD, but not UC, in the Japanese population by a candidate gene-based association study, thereby supporting the meta-analysis of the GWA data in populations of European and North American ancestry [9].

The presence of the C allele and its homozygous C/C genotype at rs2293152 SNP in *STAT3* conferred susceptibility to CD. CD- and UC-susceptible rs744166 SNP, which was identified by GWA studies [9, 12, 14], was not analyzed in this study because this SNP was not selected as a genotyped tag SNP by the iHap software program. Because rs744166 and rs957970 SNPs are located within the same linkage disequilibrium block (Fig. 1) and rs957970 SNP was not associated with susceptibility to

**Table V** Distributions of Diploypes of *TYK2* Between CD Patients and Control Subjects

Diplotype	Number (%) of diploypes in		Diplotype comparison <sup>a</sup>	
	CD	Control	OR (95% CI)	<i>P</i> value
Hap 1/Hap 1	37 (44.6)	61 (30.5)	1.833 (1.082–3.105)	0.024
Hap 1/Hap 2	20 (24.1)	73 (36.5)	0.552 (0.309–0.986)	0.045
Hap 1/Hap 3	15 (18.1)	24 (12.0)	1.618 (0.801–3.268)	0.180
Hap 1/Hap 4	0	2 (1.0)	–	–
Hap 2/Hap 2	6 (7.2)	23 (11.5)	0.600 (0.235–1.531)	0.285
Hap 2/Hap 3	5 (6.0)	15 (7.5)	0.791 (0.278–2.251)	0.660
Hap 2/Hap 4	0	1 (0.5)	–	–
Hap 3/Hap 3	0	1 (0.5)	–	–
Total number	83	200		

OR odds ratio, CI confidence interval

<sup>a</sup> Each diplotype was compared to other diploypes combined using a logistic regression analysis



**Table VI** Gene–Gene Interaction Between *STAT3* Genotype and *TYK2* Diplotype for Susceptibility to CD

Factor	Factor comparison <sup>a</sup>	
	OR (95% CI)	P value
C/C genotype at rs2293152 in <i>STAT3</i>	3.113 (1.515–6.399)	0.002
Hap 1/Hap 1 diplotype of <i>TYK2</i>	1.783 (1.042–3.053)	0.030

OR odds ratio, CI confidence interval

<sup>a</sup> Factors were statistically analyzed by a multivariate logistic regression analysis

CD in this study, rs744166 SNP may not be associated with CD in the Japanese population. The difference in susceptibility to CD at the SNP site between Caucasian and Japanese subjects can be attributed to genetic background, although *STAT3* may contribute to the same mechanisms of the immunopathogenesis of CD in both Caucasian and Japanese patients.

The presence of the A allele and its homozygous A/A genotype at rs280519 SNP in *TYK2*, the C allele and its homozygous C/C genotype at rs2304256 SNP in *TYK2*, the Hap 1 haplotype (A allele at rs280519 SNP and C allele at rs2304256 SNP) of *TYK2*, and its homozygous Hap 1/Hap 1 diplotype of *TYK2* showed susceptibility to CD. CD-susceptible rs12720356 SNP, which was identified by GWA studies [10], was not analyzed in this study because this SNP was not selected as a genotyped tag SNP by the iHap software program. Although rs280519 and rs2304256 SNPs, which were examined in this study, are located within the same linkage disequilibrium block, rs12720356 SNP does not belong to any linkage disequilibrium blocks (Fig. 2). Furthermore, because rs280496 SNP, nearby rs12720356 SNP, was not associated with susceptibility to CD in this study, rs12720356 SNP may thus not be associated with CD in the Japanese population. This disparity can be also attributed to genetic differences between Caucasian and Japanese individuals, although *TYK2* may contribute to the same immunopathogenesis in both Caucasian and Japanese CD patients.

In addition, the presence of both the C/C genotype at rs2293152 in *STAT3* and the Hap 1/Hap 1 diplotype of

*TYK2* independently contributed to the pathogenesis of CD and remarkably increased the odds ratio for CD, thus indicating an approximately 7.5-fold increase in susceptibility to CD in this study, although such CD patients account for only approximately 13% (11 of 83=13.3% in Table VII) of the genetic variance observed in CD. These findings imply that *STAT3* and *TYK2* are genetic determinants for the predisposition to the onset and/or development of CD in Japanese individuals. However, this study population was relatively small, and further studies on a larger number of Japanese subjects and on other ethnicities are necessary to confirm the association between the *STAT3* and *TYK2* polymorphisms and CD. Additional studies are needed because different populations will often have different allele frequencies and haplotype structures.

Recent GWA studies on the *IL-23/IL23R* signaling pathway have shifted the focus to the IL-23 cytokine [6–12]. After IL-23 binds to the receptor, which comprises IL23R and IL12RB1 [35], IL-23 signaling may induce the activation of JAK2 in IL23R as well as TYK2 in IL12RB1 because the IL-12RB1 and IL-23R require TYK2 [16], thus resulting in the phosphorylation of STAT3 as well as STAT1, STAT4, and STAT5 in activated macrophages and dendritic cells [35]. The signaling cascade eventually leads to the differentiation of CD4<sup>+</sup> (naive) T cells into Th17 cells [17, 20–22]. Th17 cells produce IL-17A, IL-17F, and IL-22, which are involved in the first line of the host defense by controlling the immune responses [22]. Indeed, the expression of IL-12, IL-23, STAT3, IL-17, and IL-22 has been reported to increase in the lamina propria of the

**Table VII** The Gene–Gene Combination Effect of *STAT3* Genotype and *TYK2* Diplotype for Susceptibility to CD

Factor	Number (%) of		Factor comparison <sup>a</sup>	
	CD	Control	OR (95% CI)	P value
C/C genotype at rs2293152 in <i>STAT3</i> and Hap 1/Hap 1 of <i>TYK2</i>	11 (13.3)	4 (2.0)	7.486 (2.310–24.261)	0.0008
Other genotypes	72 (86.7)	196 (98.0)		

OR odds ratio, CI confidence interval

<sup>a</sup> Factors were statistically analyzed by a multivariate logistic regression analysis

intestinal mucosa in CD patients [36–41]. Taken together, the IL-23/IL23R signaling pathway is central to the inflammation leading to CD and modifies an individual's risk of developing CD. For these reasons, it may be speculated that the polymorphisms of *STAT3* and *TYK2*, especially the C/C genotype at rs2293152 in *STAT3* and the Hap 1/Hap 1 diplotype of *TYK2*, may affect the gain-of-function of both *STAT3* and *TYK2*, thus altering the efficiency of the IL-23/IL23R signaling pathway. These changes can lead to the perpetuation of the chronic intestinal inflammatory process, thereby resulting in the onset and/or development of CD.

## Conclusions

As *TYK2* and *STAT3* appear to be the genetic determinants of CD in the Japanese population, the combination polymorphism of *TYK2* and *STAT3* may be useful as a new DNA-based diagnostic biomarker for identifying high-risk individuals susceptible to CD. Finally, *STAT3* and *TYK2* may be good target molecules for the development of novel drugs in the future.

**Acknowledgments** We are grateful to the physicians, patients, and volunteers who participated in this study. This work was supported by a research grant from the Non Profit Organization Aimed to Support Community Medicine Research in Nagasaki.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Fiocchi C. Inflammatory bowel disease: etiology and pathogenesis. *Gastroenterology*. 1998;115:182–205.
2. Podolsky DK. Inflammatory bowel disease. *N Engl J Med*. 2002;347:417–29.
3. Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature*. 2007;448:427–34.
4. Cho JH. The genetics and immunopathogenesis of inflammatory bowel disease. *Nat Rev Immunol*. 2008;8:458–66.
5. Budarf ML, Labbé C, David G, Rioux JD. GWA studies: rewriting the story of IBD. *Trend Genet*. 2009;25:137–46.
6. Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, et al. A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene. *Science*. 2006;314:1461–3.
7. Rioux JD, Xavier RJ, Taylor KD, Silverberg MS, Goyette P, Huett A, et al. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet*. 2007;39:596–604.
8. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14, 000 cases of seven common diseases and 3, 000 shared controls. *Nature*. 2007;447:661–78.
9. Cardon LR, Anderson CA, Drummond H, Nimmo E, Ahmad T, Prescott NJ, et al. Genome-wide association study defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet*. 2008;40:955–62.
10. Wang K, Zhang H, Kugathasan S, Annesse V, Bradfield JP, Russell RK, et al. Diverse genome-wide association studies associate the *IL12/IL23* pathway with Crohn disease. *Am J Hum Genet*. 2009;84:399–405.
11. Lobo A, Forbes A, Sanderson J, Jewell DP, Mansfield JC, Deloukas P, et al. Genetic determinants of ulcerative colitis induce the *ECMI* locus and five loci implicated in Crohn's disease. *Nat Genet*. 2008;40:710–2.
12. Franke A, Balschun T, Karlsen TH, Hedderich J, May S, Lu T, et al. Replication of signals from recent studies of Crohn's disease identifies previously unknown disease loci for ulcerative colitis. *Nat Genet*. 2008;40:713–5.
13. Abraham C, Cho JH. IL-23 and autoimmunity: new insights into the pathogenesis of inflammatory bowel disease. *Annu Rev Med*. 2009;60:97–110.
14. Mathew CG, Schreiber S, the IBSEN study group, Franke A, Balschun T, Karlsen TH, et al. Sequence variants in *IL10*, *APRC2* and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet*. 2008;40:1319–23.
15. Shimoda K, Kato K, Aoki K, Matsuda T, Miyamoto A, Shibamori M, et al. Tyk2 plays a restricted role in IFN $\alpha$  signaling, although it is required for IL-12-mediated T cell function. *Immunity*. 2000;13:561–71.
16. Murray PJ. The JAK-STAT signaling pathways: input and output integration. *J Immunol*. 2007;178:2623–9.
17. Watford WT, Hissong BD, Bream JH, Kanno Y, Muul L, O'Shea JJ. Signaling by IL-12 and IL-23 and the immunoregulatory roles of *STAT4*. *Immunol Rev*. 2004;202:139–56.
18. Heinrich PC, Behrmann I, Müller-Newen G, Schaper F, Graeve L. Interleukin-6-type cytokine signalling through the gp130/Jak/STAT pathway. *Biochem J*. 1998;334:297–314.
19. Minegishi Y, Saito M, Morio T, Watanabe K, Agematsu K, Tsuchiya S, et al. Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity. *Immunity*. 2006;25:745–55.
20. Kikly K, Liu L, Na S, Sedgwick JD. The IL-23/Th<sub>17</sub> axis: therapeutic targets for autoimmune inflammation. *Curr Opin Immunol*. 2006;18:670–5.
21. Rachitskaya AV, Hansen AM, Horai R, Li Z, Villasmil R, Luger D, et al. Cutting edge: NKT cells constitutively express IL-23 receptor and ROR $\gamma$ t and rapidly produce IL-17 upon receptor ligation in an IL-6-independent fashion. *J Immunol*. 2008;180:5167–71.
22. Nakamura R, Shibata K, Yamada H, Shimoda K, Nakayama K, Yoshikai Y. Tyk2-signaling plays an important role in host defense against *Escherichia coli* through IL-23-induced IL-17 production by  $\gamma\delta$  T cells. *J Immunol*. 2008;181:2071–5.
23. Lennard-Jones JE. Classification of inflammatory bowel disease. *Scand J Gastroenterol Suppl*. 1989;170:2–6.
24. Podolsky DK. Inflammatory bowel disease (1). *N Eng J Med*. 1991;325:928–37.
25. Podolsky DK. Inflammatory bowel disease (2). *N Eng J Med*. 1991;325:1008–16.
26. Satsangi J, Silverberg MS, Vermeire S, Colombel JF. The Montreal classification of inflammatory bowel disease: controversies, consensus, and implications. *Gut*. 2006;55:749–53.
27. Rachmilewitz D. Coated mesalazine (5-aminosalicylic acid) versus sulphasalazine in the treatment of active ulcerative colitis: a randomised trial. *BMJ*. 1989;298:82–6.
28. Best WR, Beckett JM, Singleton JW, Kern F Jr. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology*. 1976;70:439–44.
29. Akira S, Nishio Y, Inoue M, Wang X-J, Wei S, Matsusaki T, et al. Molecular cloning of APRF, a novel IFN-stimulated gene factor 3

- p91-related transcription factor involved in the gp130-mediated signaling pathway. *Cell*. 1994;77:63–71.
30. Firmbach-Kraft I, Byers M, Shows T, Dalla-Favera R, Krolewski JJ. tyk2, prototype of a novel class of non-receptor tyrosine kinase genes. *Oncogene*. 1990;5:1329–36.
  31. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007; 449:851–61.
  32. Song CM, Yeo BH, Tantoso E, Yang Y, Lim YP, Li K-B, et al. iHAP—integrated haplotype analysis pipeline for characterizing the haplotype structure of genes. *BMC Bioinformatics*. 2006;7:525.
  33. Daly MJ, Rioux JD, Schaffner SF, Hudson TJ, Lander ES. High-resolution haplotype structure in the human genome. *Nat Genet*. 2001;29:229–32.
  34. Johnson GCL, Esposito L, Barratt BJ, Smith AN, Heward J, Di Genova G, et al. Haplotype tagging for the identification of common disease genes. *Nat Genet*. 2001;29:233–7.
  35. Parham C, Chirica M, Timans J, Vaisberg E, Travis M, Cheung J, et al. A receptor for the heterodimeric cytokine IL-23 is composed of IL-12R $\beta$ 1 and a novel cytokine receptor subunit, IL-23R. *J Immunol*. 2002;168:5699–708.
  36. Atreya R, Mudter J, Finotto S, Müllberg J, Jostock T, Wirtz S, et al. Blockade of interleukin 6 *trans* signaling suppresses T-cell resistance against apoptosis in chronic intestinal inflammation: evidence in Crohn disease and experimental colitis *in vivo*. *Nat Med*. 2000;6:583–8.
  37. Schmidt C, Giese T, Ludwig B, Mueller-Molaian I, Marth T, Zeuzem S, et al. Expression of interleukin-12-related cytokine transcripts in inflammatory bowel disease: elevated interleukin-23p19 and interleukin-27p28 in Crohn's disease but not in ulcerative colitis. *Inflamm Bowel Dis*. 2005;11:16–23.
  38. Mudter J, Weigmann B, Bartsch B, Kiesslich R, Strand D, Galle PR, et al. Activation pattern of signal transducers and activators of transcription (STAT) factors in inflammatory bowel diseases. *Am J Gastroenterol*. 2005;100:64–72.
  39. Fuss IJ, Becker C, Yang Z, Groden C, Hornung RL, Heller F, et al. Both iL-12p70 and IL-23 are synthesized during active Crohn's disease and are down-regulated by treatment with anti-IL-12p40 monoclonal antibody. *Inflamm Bowel Dis*. 2006;12:9–15.
  40. Wolk K, Witte E, Hoffmann U, Doecke W-D, Endesfelder S, Asadullah K, et al. IL-22 induces lipopolysaccharide-binding protein in hepatocytes: a potential systemic role of IL-22 in Crohn's disease. *J Immunol*. 2007;178:5973–81.
  41. Hölttä V, Klemetti P, Sipponen T, Westerholm-Ormio M, Kociubinski G, Salo H, et al. IL-23/IL-17 immunity as a hallmark of Crohn's disease. *Inflamm Bowel Dis*. 2008;14:1175–84.