Identification of germline variation associated with thiopurine-induced leukopenia in inflammatory bowel disease

Han Sang Kim

Department of Medical Science

The Graduate School, Yonsei University

Identification of germline variation associated with thiopurine-induced leukopenia in inflammatory bowel disease

Han Sang Kim

Department of Medical Science

The Graduate School, Yonsei University

Identification of germline variation associated with thiopurine-induced leukopenia in inflammatory bowel disease

Directed by Professor Min Goo Lee

The Doctoral Dissertation
submitted to the Department of Medical Science,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Han Sang Kim

December 2017

This certifies that the Doctoral Dissertation of Han Sang Kim is approved.

Thesis Supervisor : Min Goo Lee

Thesis Committee Member#1 : Sun Young Rha

Thesis Committee Member#2 : Chul Hoon Kim

Thesis Committee Member#2 : Wan Namkung

Thesis Committee Member#4 : Heon Yung Gee

The Graduate School
Yonsei University

December 2017

ACKNOWLEDGEMENTS

내과 전문의로서 의과학자의 꿈을 가지고 31살의 조금 늦은 나이로 도전한 5년의 박사 과정이 이제 마무리되려 합니다. 무엇보다 지난 5년의 시간 동안 항상 크고 든든한 버팀목이 되어 주신 이민구 지도교수님께 깊은 감사의 말씀을 드립니다. 기초 의과학에 대한 뛰어난 지식뿐만 아니라 제자들이 잘 될 수 있도록 항상 배려해 주시고 노력해 주신 참 스승님의 따뜻한 모습을 잊을 수가 없습니다. 매번 어려움을 겪을 때마다 격려해 주시고 한 발자국 나아갈 수 있도록 격려해 주셔서 감사드리며 앞으로 가야할 길이 더 멀지만 그 격려의 말씀을 잊지 않고 계속 발전할 수 있도록 노력하겠습니다.

박사 학위를 지도해주신 심사위원 교수님들께도 깊은 감사의 말씀을 올립니다. 15년전 의과대학 학생때부터 시작하여 석사학위를 지도해 주시면서 중개의학의 중요성을 가르쳐 주시고 종양내과를 전공할 수 있도록 이끌어 주신 라선영 교수님, 항상 교수실의 문을 활짝 열어놓고 여러 분야의 전문가들과 소통하시는 김철훈 교수님, 연구실의 가장 1대 선배님 이면서 후배들의 귀감이 되어 주시는 남궁완 교수님, 누구보다 먼저 실험실에 나와 실험하시고 주말에도 연구에 대한 열정과 타의 모범을 보여주신 지헌영 교수님께 깊은 감사의 말씀을 올립니다. 모든 선생님들께서 지도해 주신 덕분으로 많은 것을 배우고 느끼며 성장할 수 있었습니다.

또한 박사 학위를 받는 동안 귀감이 되어 주신 여러 교수님들이 있기에 더 많이 성장할 수 있었습니다. 항상 새로운 시도와 연세의료원의 연구발전을 위해 노력하시는 백순명 교수님, 유전체 공부를 하면서 한 단계 성장할 수 있도록 도와 주신 김태민 교수님, 석사때부터 이어져온 인연으로 항상 여러가지로 도와주시고 국가생명연구자원 정보센터에서 잊지 못할 소중한 경험을 하도록 도움 주신 김상철 박사님께도 감사의 말씀을 올립니다. 소중한 공동 연구의 기회를 주신 방두희 교수님, 이지현 교수님, 김상우 교수님, 김형범 교수님께도 감사의 말씀을 드립니다.

종양내과 의사로서 성장하는데 조언을 아끼지 않으시는 종양내과

정현철 교수님, 안중배 교수님, 손주혁 교수님, 조병철 교수님, 최혜진 교수님, 신상준 교수님, 정민규 교수님, 김혜련 교수님, 김효송 교수님, 김건민 교수님께도 감사의 말씀을 드립니다.

종양 생물학 공부를 같이하며 박사 학위 하는 동안 여러 기쁨과 어려움을 함께 나눈 임상의과학자 선배님들께도 감사의 말씀을 전하고 싶습니다. 항상 조언해 주시는 윤홍인 선배님, 같은 종양내과를 전공하며 친형처럼 아껴 주시는 박형순 선배님, 치과대학의 차용훈 선배님께도 감사의 말씀을 드립니다. 고등학교 때부터 이어져 항상 여러가지 도움 주고 매번 조언을 아끼지 않는 김태균, 그리고 우리 임상 의과학자 프로그램 동기 조용근, 한현호, 이승현 모두 고맙고 앞으로도 힘내자.

연구실에서 항상 어려운 일이 있을 때마다 힘이 되어준 김연정 선생님, 채동우 형, 실험실 동기면서 동고동락을 같이한 김지윤, 윤종진, 국진주 선생님, 노신혜 누나, 엄소원, 박학, 정우영 선생님, 허운, 고영익, 조경지, 이준석과 같은 유전체 파트에서 공부한 박준희, 한수민, 윤지훈, 지금은 내과 후배이기도 한 은성호에게도 감사의 말씀을 드립니다.

Especially, I would like to express my deep and sincere gratitude to my mentor, Professor David Lyden, M.D., Ph.D., a father of exosome and pre-metastatic niche, whose stimulating suggestions and encouragement helped me throughput my doctoral degree. As a pioneer and a visionary physician-scientist, he guided me into the right path to become a translational researcher, still a long road ahead. Whenever I need advice, he gladly helped me a lot to grow up in spite of his crowded schedule. With his suggestions and encouragement, I would like to build my successful career as translational researcher.

마지막으로 항상 든든한 가족이 있기에 어려운 공부를 할 수 있었던 것 같습니다. 양가 부모님, 누나 매형, 처남 처남댁, 조카들 그리고 항상 저를 이해해주고 응원해 주는 아내 홍선이와 두 딸 도유, 도아 모두 고맙고 사랑합니다.

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Identification of germline variation associated with thiopurine-induced leukopenia in inflammatory bowel disease

Han Sang Kim

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Min Goo Lee)

Myelosuppression is a life-threatening complication of thiopurine therapy, and the incidence of thiopurine-induced myelosuppression is higher in East Asians than in Europeans. We investigated genetic factors associated with thiopurine-induced leukopenia in patients with inflammatory bowel disease (IBD). A genome-wide association study (GWAS) was conducted in thiopurine-treated patients with IBD, followed by high-throughput sequencing of genes identified as significant in the GWAS or those involved in thiopurine metabolism (n = 331). Significant loci associated with thiopurine-induced leukopenia were validated in two additional replication cohorts (n = 437 and n = 330). Functional consequences of FTO (fat mass and obesity-associated) variant were examined both in vitro and in vivo. The GWAS identified two loci associated with thiopurine-induced leukopenia (rs16957920, FTO intron; rs2834826, RUNX1 intergenic). High-throughput targeted sequencing indicated that an FTO coding variant (rs79206939, p.A134T) linked to rs16957920 is associated with thiopurine-induced leukopenia. This result was further validated in two replication cohorts (combined P = 1.3×10-8, odds ratio = 4.3). The frequency of FTO A134T is 5.1% in Koreans but less than 0.1% in Western populations. The p.A134T variation reduced FTO activity by 65% in the nucleotide demethylase assay. In vivo experiments revealed that Fto−/− and

Fto+/− mice were more susceptible to thiopurine-induced myelosuppression than wild-type mice. The results suggest that the hypomorphic FTO A134T variant is associated with thiopurine-induced leukopenia. These results not only shed light on the novel physiological role of FTO but also provide a potential pharmacogenetic biomarker for thiopurine therapy.

Key words: GWAS, azathioprine, leukopenia, inflammatory bowel disease, FTO

Identification of germline variation associated with thiopurine-induced leukopenia in inflammatory bowel disease

Han Sang Kim

Department of Medical Science
The Graduate School, Yonsei University

(Directed by Professor Min Goo Lee)

I. INTRODUCTION

The thiopurine 6-mercaptopurine (6-MP) and its prodrug azathioprine (AZA) have been widely used in anti-inflammatory, anticancer, and immunosuppressive therapies for over 50 years.[1] AZA and 6-MP remain the cornerstones for inducing and maintaining remission in patients with inflammatory bowel disease (IBD).[2-4] However, over 20% of patients have to discontinue thiopurines due to drug-related adverse events.[5,6] In particular, the life-threatening bone marrow suppression is the most common reason for treatment interruption.[7,8]

The clinical characteristics of thiopurine-induced myelosuppression have been extensively investigated over the past few decades. The important clinical characteristics of thiopurine-induced myelosuppression can be summarized as follows. First, myelosuppression is a dose-dependent adverse event. Leukopenia, generally defined as a white blood cell (WBC) count of $\leq$ 3-4 $\times$ 10^9/L, is the most common indicator of myelosuppression.[9] Second, although most cases of severe myelosuppression occur early in treatment, the timing of thiopurine-induced myelosuppression varies from 12 days to 27 years after thiopurine treatment.[10,11] Third, myelosuppression may occur abruptly and without symptoms during thiopurine treatment; therefore, regular blood count

examinations (e.g., every 3 months) are recommended to prevent severe bone marrow toxicity.[12] Despite abundant data regarding the characteristics of thiopurine-induced myelosuppression, clinical information alone is not sufficient to anticipate and manage the risk of developing bone marrow suppression.

The therapeutic efficacy and toxicity of thiopurines are closely linked to their metabolism. Thiopurines are prodrugs that are converted into active metabolites via intricate metabolic processes, including the purine salvage pathway.[13] AZA is converted to 6-MP in erythrocytes by a non-enzymatic reaction involving glutathione. After transport into cells, 6-MP is transformed into nucleoside monophosphates (thioinosine monophosphate [TIMP]) by hypoxanthine-guanine-phosphoribosyltransferase 1 (HPRT1).[1] In addition, subsequent reactions involving kinases and reductases generate thioguanine nucleotides (TGNs) from TIMP.[1] TGNs are incorporated into DNA and RNA, thereby inhibiting cell proliferation by blocking DNA and RNA synthesis. Meanwhile in the process of catabolism, thiopurines can be inactivated by thiopurine S-methyltransferase (TPMT) or xanthine oxidase.[13] Interestingly, TPMT can also generate methyl-TIMP, which is a powerful inhibitor of de novo purine biosynthesis.[1]

To date, several studies have attempted to identify genetic determinants of thiopurine-induced toxicity. In Caucasians, genetic variations in TPMT have been identified as a major cause of thiopurine-induced myelosuppression. [14, 15] Interestingly, East Asian people are more susceptible to develop thiopurine-induce myelosuppression, although TPMT polymorphisms that induce the TPMT enzyme deficiency are rare in East Asians.[14] The frequency of life-threatening bone marrow suppression is 1.4–5% in Western countries and 31–56% in East Asian countries.[7,8] A recent report suggested that NUDT15 R139C is strongly correlated with thiopurine-induced early leukopenia in

Korean patients with Crohn's disease.[15] However, a majority of thiopurine-induced leukopenia cases cannot be explained by genetic variations in TPMT or NUDT15, as many patients with wild-type TPMT and NUDT15 experience clinically significant thiopurine-induced myelosuppression.[16-18]

The aim of the present study was to identify novel genetic factors associated with thiopurine-induced leukopenia in patients with IBD. We performed a genome-wide association study (GWAS) followed by high-throughput sequencing of coding regions to comprehensively investigate variations in genes identified as significant in the GWAS, as well as genes involved in thiopurine metabolism and transport pathways. The results of the initial GWAS and high-throughput sequencing were re-evaluated through replication studies in two additional cohorts and through *in vitro* and *in vivo* molecular functional studies.

II. MATERIALS AND METHODS

Study population and design

The study design is illustrated in Figure 1. A total of 1,098 patients with IBD who had been treated with thiopurines were included in this study. The discovery GWAS cohort (n = 331) enrolled at Severance Hospital in Seoul, Korea, between July 1998 and January 2012 (CD, n=161; UC, n=107; Bechet disease, n=63). An external validation cohort (Replication cohort 1) consisted of 437 patients with Crohn's disease enrolled at the Crohn's Disease Research Network, Korea (CD, n=185),[19] and the Tohoku University Hospital, Japan, between April 2002 and September 2015 (CD, n=145; UC, n=106; Bechet disease, n=1). An internal replication cohort (Replication cohort 2) consisted of 330 patients enrolled at the Severance Hospital between December 2011 and March 2015 (CD, n=228; UC, n=102). All enrolled patients had East Asian ethnicity and there was no overlap of samples among studied cohorts. Criteria

for leukopenia included a white blood cell (WBC) count of ≤ 3,000/mm^3 and dose interruption (e.g., withdrawal or dose reduction) during thiopurine treatment. Representative examples of patients with and without leukopenia are illustrated in Figure 2. This study was approved by the Institutional Review Board of Severance Hospital (4-2011-0751).

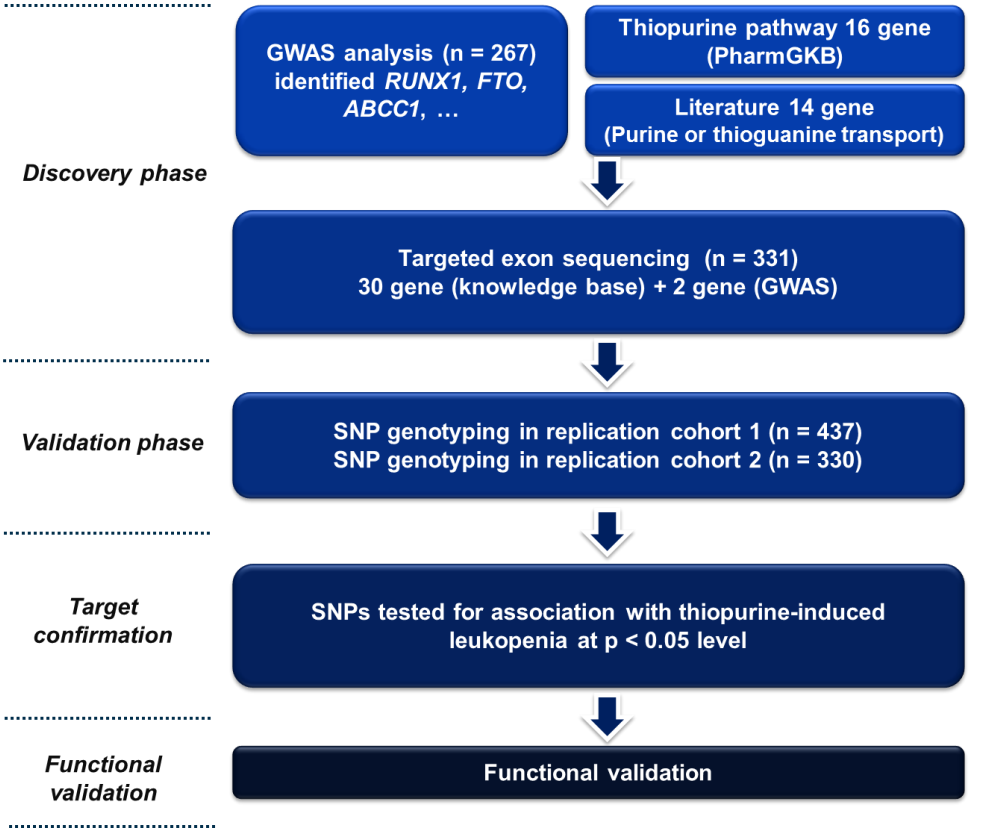


Figure 1. Study scheme.

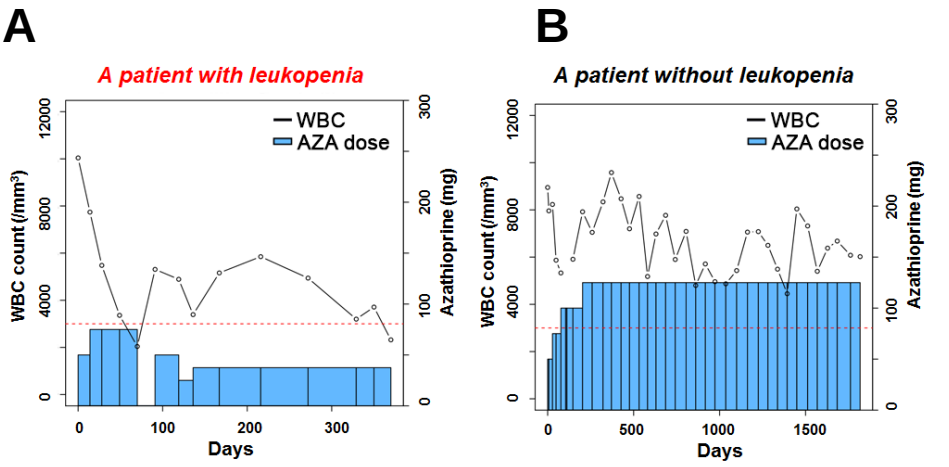


Figure 2. Representative examples of patients with and without leukopenia in the case-control study. Blue lines indicate white blood cell count (*left y-axis*), and blue bar graph indicates dose of azathioprine (*right y-axis*) by day (*x-axis*) from the day when azathioprine was first prescribed. One patient with leukopenia (*A*) had to stop taking azathioprine due to leukopenia, whereas one patient without leukopenia (*B*) used azathioprine for more than 3 years without dose interruption. Red dashed line indicates a white blood count of 3,000/mm^3.

The use of thiopurines

Physicians usually prescribed AZA at a dose of 25–50 mg (or 6-MP at a dose of 12.5–25 mg) daily first, and increased the AZA dose of 25–50 mg (or 12.5–25 mg for 6-MP) every 2–4 weeks to achieve 2.0–3.0 mg/kg/day for AZA and 1.0–1.5 mg/kg/day for 6-MP, the therapeutic doses for immunosuppression. Given that the molecular weight of AZA is 55% that of 6-MP, converting to the AZA dose can be achieved by multiplying the 6-MP dose by 2.08. A complete blood count (CBC) was performed every 1–2 weeks until the therapeutic dose was achieved, and every 2–3 months thereafter.

Genotyping for GWAS

Genome-wide genotyping was performed using an Axiom Genome-Wide ASI 1 Array Chip (Affymetrix, CA, USA), which was designed to maximize the coverage of rare variants in East Asian populations. Data quality control was

performed according to a method described previously[20,21] using PLINK software, ver1.07, as illustrated in Figure 3. Detailed information for quality conrtol is available in the Supplementary Information.

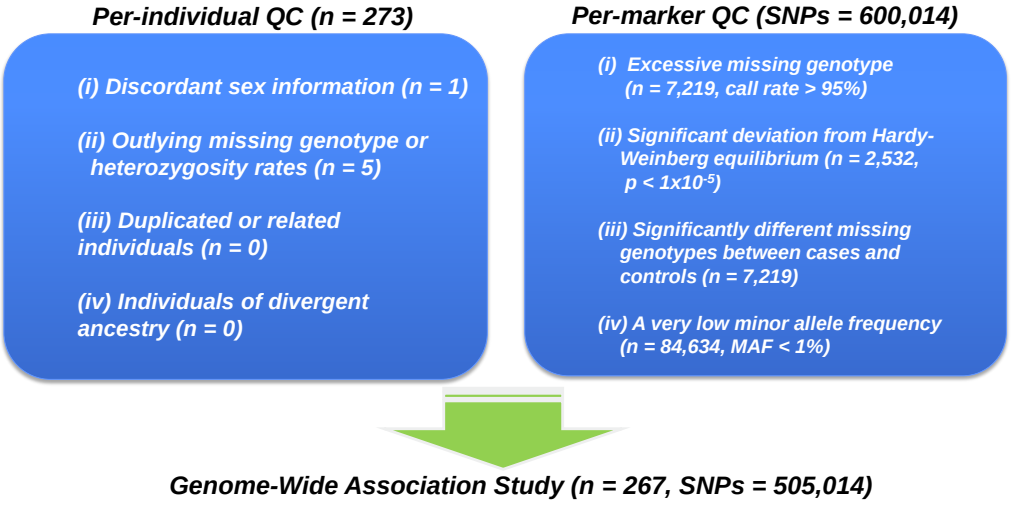


Figure 3. Data quality control (QC) in the genome-wide association study.

Quality control for GWAS

Per-individual quality control was performed to exclude individuals with discordant sex information and outlying missing genotypes, resulting in the inclusion of 267 of 273 samples. Next, per–marker quality control was performed based on call rate (call rate ≥ 95%), minor allele frequency (MAF; ≥ 1%), and Hardy-Weinberg equilibrium test ($P < 1\times10^{-5}$), resulting in the inclusion of 505,014 of 600,014 single nucleotide polymorphism (SNP) markers in the GWAS (Figure 3).

High-throughput sequencing and analysis

After the GWAS, fine mapping of the coding regions surrounding the top markers, as well as of genes previously reported as related to thiopurine

metabolism, was performed for the discovery cohort (n = 331; including 267 patients analyzed in the GWAS). A total of 32 genes, including 16 genes involved in the thiopurine metabolism pathway identified by PharmGKB,[13] 14 genes involved in purine or thioguanine transport,[22,23] and 2 genes identified in the GWAS, were included (Table 1). Target enrichment from blood samples was performed using a Human In-solution Hybrid Capture Kit (Celemics, Seoul, Korea), and paired-end sequencing (2×150 bp) was carried out using an Illumina HiSeq 2500 sequencing platform. The analysis pipeline is illustrated in Figure 4. *FTO* germline variants were investigated in exome sequencing data from a control population consisting of 842 healthy Koreans with no history of IBD (mean depth, 65X). There was no difference of genotyping results between GWAS and high-throughput sequencing. All sequencing reads were aligned with the NCBI build 37 (hg19) human reference genome using the Burrows-Wheeler Aligner.[24] PCR duplicates were removed using Picard (http://picard.sourceforge.net), and local re-aligning around indels and paired-end fixing were performed using the Genome Analysis Toolkit (GATK).[25] Variant calling was performed using GATK HaplotypeCaller,[25] and genetic variants were annotated using ANNOVAR.[26]

Table 1. List of 32 genes for high-throughput sequencing

Thiopurine pathway from PharmGKB (16 genes)	Purine or thioguanine transport from literatures (14 genes)	Candidates from GWAS (2 genes)
TPMT	*ABCB1*	*FTO*
ITPA	*ABCG2*	*RUNX1*
XDH	*ABCC2*	
HPRT1	*SLC22A1*	
IMPDH1	*SLC22A2*	
GMPS	*SLC22A3*	
GSTA1	*SLC15A1*	
GSTA2	*SLC22A6*	
GSTM1	*SLC22A7*	
AOX1	*SLC22A8*	
SLC28A2	*MTAP*	
SLC28A3	*NT5C2*	
SLC29A1	*MOCOS*	
SLC29A2	*MTHFR*	
ABCC4		
ABCC5		

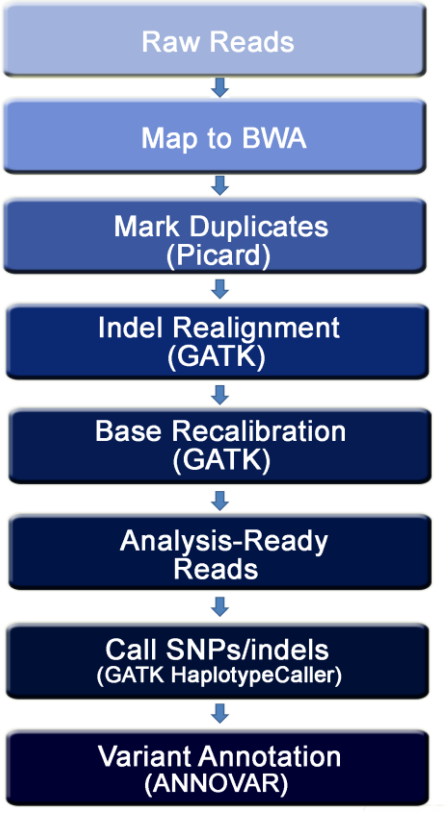


Figure 4. High-throughput sequencing analysis pipeline.

SNP genotyping

To verify significant loci in the discovery and validation cohorts, rs79206939, rs16952570, rs2834826, and rs116855232 were genotyped using the TaqMan SNP genotyping assay (Thermo Fisher Scientific, MA, USA) for the discovery cohort and Replication cohort 1. For SNP genotyping in Replication cohort 2, rs16952570, rs2834826, and rs116855232 were genotyped by the SNaPshot assay (Thermo Fisher Scientific, MA, USA), and rs79206939 was genotyped by a Sanger sequencing.

Cell culture and immunoblotting

Jurkat immortalized human T lymphocytes were maintained in RPMI 1640 medium (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum and penicillin (50 IU/mL)/streptomycin (50 μg/mL). The expression vector for the coding region of human *FTO* (pCMV3-hFTO; c-terminal myc-tagged; GenBank NM_001080432.2) was purchased from Sino Bio (Beijing, China). *FTO* c.400G>A (*FTO* p.A134T) mutant expression plasmids were generated with a PCR-based site-directed mutagenesis kit (Stratagene, CA, USA). Jurkat cells stably expressing FTO wild type (FTO^{WT}) or FTO mutant type (FTO^{A134T}) were constructed through lentiviral transduction. Immunoblotting was performed as described previously.[27] Rat polyclonal antibody against puromycin N-acetyltransferase (PAC) was generated using recombinant His6-tagged PAC, and antibodies against c-Myc (sc-40; Santa Cruz Biotechnology, TX, USA) and ß-actin (M5546; Sigma, MO, USA) were from commercial sources.

FTO 3-Methyluracil demethylation assay

The fluorescent FTO demethylation assay was performed as described elsewhere.[28] Briefly, recombinant wild-type and mutant FTO (FTO^{A134T} and FTO^{R316Q}) proteins were purified, and the stem-loop substrate containing a methylated uridine at the N-3 position was manufactured. In each reaction, methylated substrate (100 nM) was incubated with 75 μM $Fe(NH_4)_2(SO_4)_2$, 300 μM 2-oxoglutarate, 2 mM ascorbate, 50 μg/mL bovine serum albumin, 625 pg of RNase A, and various concentrations of wild-type or mutant FTO at a concentration ranging from 0 to 1,000 nM. 6-Carboxyfluorescein emission was measured for 30 min at a wavelength of 535 nm (with excitation at 485 nm) using a Tecan Infinite f500 microplate reader in a dark flat-bottomed 96-well plate at room temperature.

***In vivo* thiopurine treatment studies**

The generation and characterization of Fto mutant mice (C57BL/6J) were described previously.[29] All experimental protocols were approved by the Institutional Animal Care and Use Committee, Yonsei University College of Medicine (protocol number 2014-0175).

Thiopurines (AZA or 6-MP) were suspended in phosphate-buffered saline (PBS) and administered by gavage (AZA, 30 mg/kg) or intraperitoneal injection (6-MP, 10 mg/kg) in a volume of approximately 0.1–0.2 mL/mouse. Control mice were treated with vehicle (PBS) at the same dose volume. Blood sampling from the retro-orbital plexus was performed on day 0 (baseline) and day 15 after thiopurine treatment. Hematologic traits were measured using a Hemavet 950 (Drew Scientific, FL, USA). The femur was removed and placed in 4% phosphate-buffered formalin fixative for bone marrow examination. Tissues were embedded in paraffin, sectioned, and then stained with hematoxylin and eosin for morphologic examination using light microscopy. Bone marrow cells were counted using ImageJ (ver. 1.49).

Statistical analysis

The GWAS was performed using the allelic chi-square and Wald tests implemented in PLINK software for the case-control study.[21] The SNPs showing the strongest statistical significance ($P < 5\times10^{-6}$) were selected for replication. The criteria for replication included significance at the nominal level ($P < 0.05$) and same direction of effect as observed in the GWAS. The Breslow-Day test was used to evaluate the heterogeneity among the studies. Associations between the combined subsets of samples were analyzed using the Cochran-Mantel-Haenszel test. Estimation of pair-wise linkage disequilibrium (LD) was performed using Haploview 4.2.

Results are presented as the mean ± SEM for the indicated number of in

vitro and in vivo experiments. Statistical analyses were performed using Student's t-test, log-rank test, or Kaplan-Meier method, as appropriate. A *P* value < 0.05 was considered statistically significant. Statistical analyses were conducted using R software (v3.1.0).

III. RESULTS

Genome-Wide Association Study

With the aim of discovering novel variants that confer sensitivity to thiopurines, we carried out a GWAS in the discovery cohort (n = 267). The baseline clinical characteristics of the discovery cohort are shown in Table 2. The enrolled IBD patients were diagnosed with Crohn's disease (59%), ulcerative colitis (26%), or Bechet's disease (15%). The median age at start of thiopurine therapy was 32 years and the average initial AZA dose was 1 mg/kg. The incidence of leukopenia was 42.3% (113 leukopenia and 154 no leukopenia). There was no significant difference in baseline WBC count between the leukopenia subjects and no leukopenia subjects.

As detailed in Figure 3, we performed the GWAS using 505,014 markers after applying quality control for SNPs. The genomic inflation factor (λ) was 1.00078, and a quantile-quantile plot of *P* values showed no systematic deviation from the expected null distribution (Figure 5). The GWAS based on allelic chi-square analyses identified 10 SNPs reaching the significance threshold ($P < 5\times10^{-6}$; Figure 6A, Table 3). Because two SNPs in 2p24, two SNPs in 16q12, and five SNPs in 21q22 form an LD block (Figure 7), four independent loci remain significant after an LD-based pruning on the GWAS dataset. The GWAS based on a logistic regression analysis showed the comparable results (Figure 8). Among them two loci, a proximal region of *RUNX1* (rs2834826, $P = 5.8\times10^{-8}$) and an intronic region of *FTO* (rs16952570, $P = 1.3\times10^{-6}$) showed the lowest association *P* values, and thus were subject to further analysis. Because two other intergenic loci (rs1969003 and rs7605946)

are located far distant from the nearest genes (rs1969003, 472 Kb apart from *GADL1* and 167 Kb apart from *STT3B*; rs7605946, 1 Mb apart from *LOC645949* and 563 Kb apart from *KLHL29*, respectively), they were excluded from high-throughput sequencing on the protein coding regions.

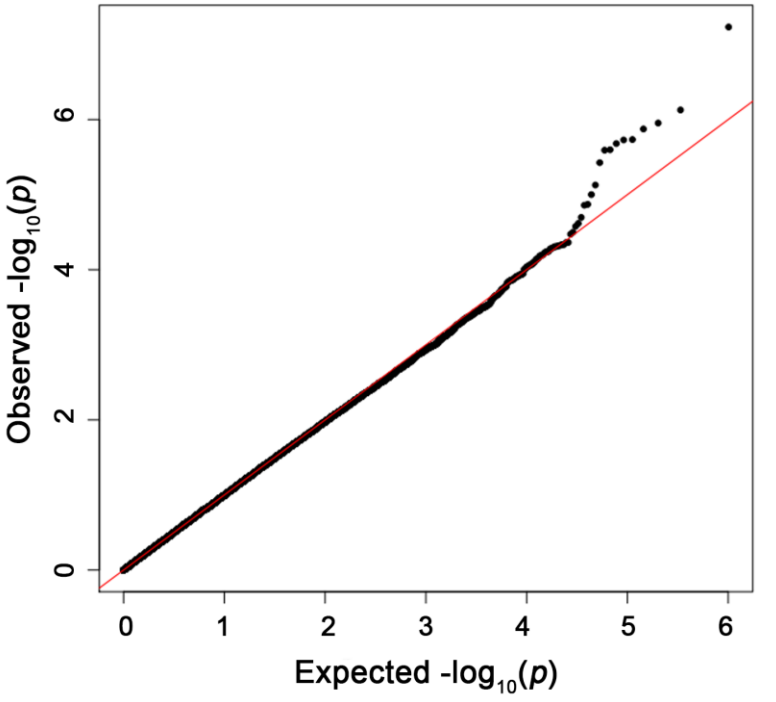


Figure 5. Quantile-quantile plot of the discovery genome-wide association study (113 leukopenia and 154 no leukopenia). Expected $-\log_{10} P$ values are those expected under the null hypothesis. Observed $-\log_{10} P$ values are the GWAS association results derived by allelic chi-square test, as shown in Figure 6A.

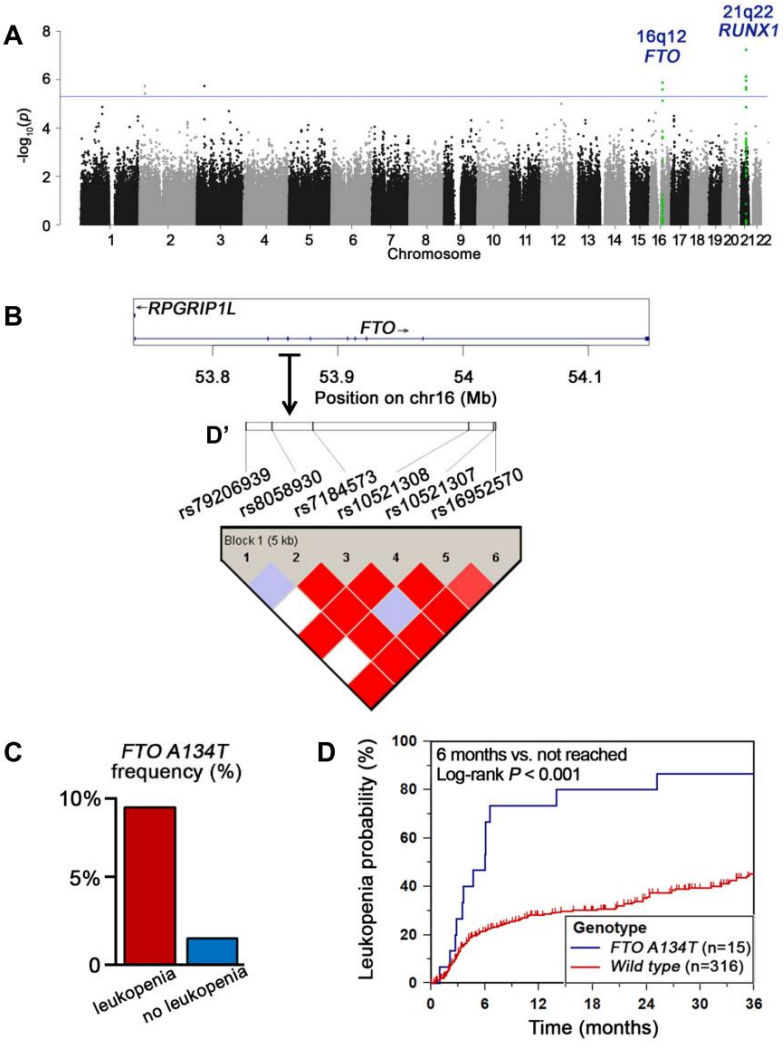


Figure 6. Genome-wide association study (GWAS) and frequency of *FTO* p.A134T in the discovery cohort. (A) Manhattan plot showing associations between SNPs and thiopurine-induced leukopenia (n = 267; 113 leukopenia and 154 no leukopenia). (B) Linkage disequilibrium (LD) in *FTO*. The linkage disequilibrium heat map indicates that *FTO* p.A134T (rs79206939, Block 1) and the *FTO* intron 3 variant (rs16952570, Block 6) are located within the same LD block (D' = 1). The standard color scheme of Haploview was applied for LD color display; logarithm of odds (LOD) score ≥2 and D′ = 1 is shown in red. (C) Genotype frequency of the *FTO* variant encoding p.A134T in patients with and without leukopenia (n = 331; 132 leukopenia and 199 no leukopenia). (D) Kaplan-Meier estimates of time-to-leukopenia between *FTO* p.A134T (n = 15) and *FTO* wild type (n = 316) in IBD patients treated with thiopurines.

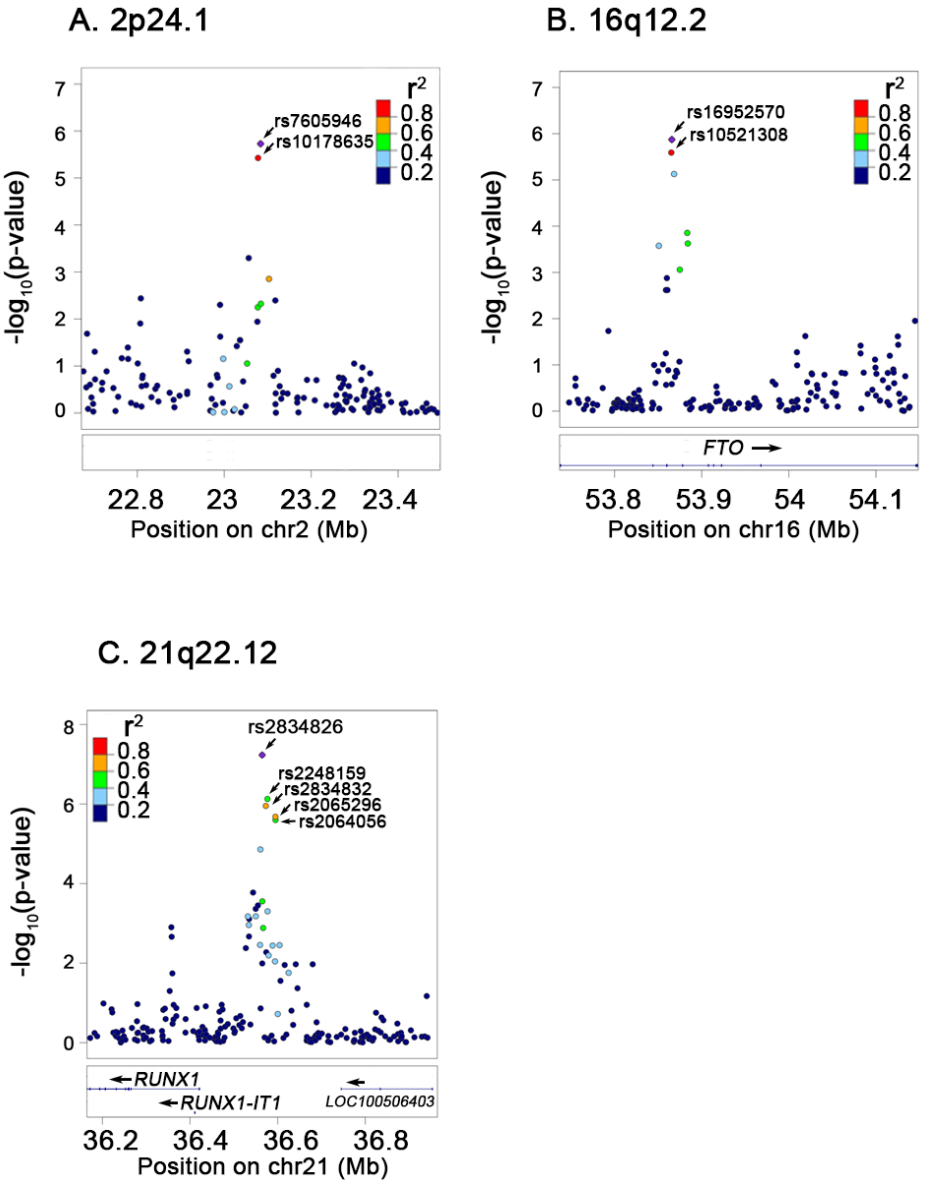
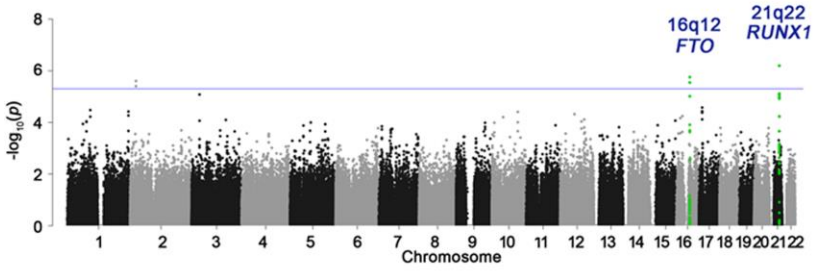


Figure 7. Regional association plot of markers in 2p24, 16q12, and 21q22 in the discovery GWAS (113 leukopenia and 154 no leukopenia).

A



B

SNP	Genomic position (bp)	Nearby gene (distance, bp)	Location	Minor allele	MAF		*P* value[a]	OR[b] (95% CI)
					Leukopenia	No leukopenia		
rs2834826	21q22.12 (36564651)	*RUNX1* (143056), *LOC100506403* (180154)	intergenic	T	56.7%	33.1%	6.3×10^{-7}	2.5 (1.8-3.6)
rs16952570	16q12.2 (53865730)	*FTO*	intronic	C	41.2%	21.8%	1.8×10^{-5}	2.8 (1.8-4.3)
rs7605946	2p24.1 (23044808)	*LOC645949* (1111284), *KLHL29* (563490)	intergenic	T	57.1%	36.3%	2.5×10^{-5}	2.6 (1.7-3.8)
rs10521308	16q12.2 (53865126)	*FTO*	intronic	A	40.6%	21.8%	2.9×10^{-5}	2.8 (1.8-4.2)
rs10178635	2p24.1 (23040566)	*LOC645949* (1107042), *KLHL29* (567732)	intergenic	T	57.5%	37.3%	4.0×10^{-5}	2.5 (1.7-3.7)

SNP, single nucleotide polymorphism; OR, odds ratio; MAF, minor allele frequency
[a,b]Results from the logistic-regression.

C

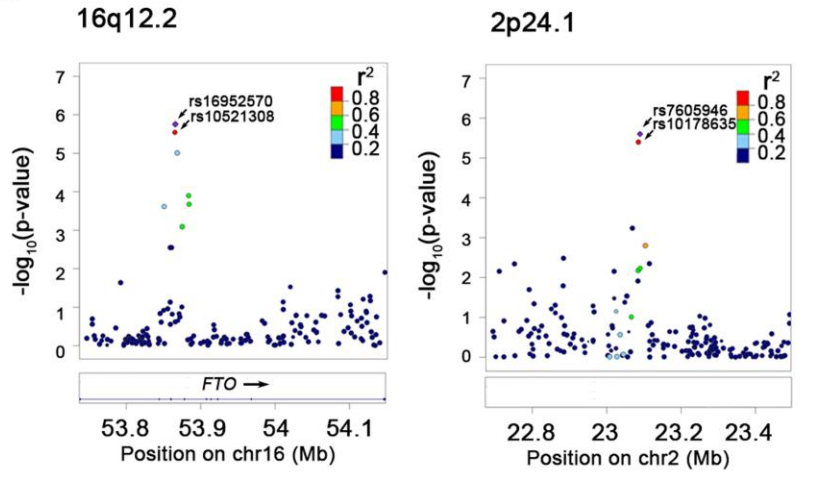


Figure 8. A logistic regression analysis in the discovery GWAS (113 leukopenia samples and 154 no leukopenia).

Table 2. Baseline clinical characteristics of the discovery cohort in genome-wide association study (n = 267)

	All patients (n = 267)	Patients without leukopenia (n = 154)	Patients with leukopenia (n = 113)	*P* value (with leukopenia *vs* without leukopenia)
Age at start of therapy, median (range)	32 (14-71)	31 (14-70)	33 (15-71)	0.16
Female sex, n (%)	86 (32)	38 (25)	48 (42)	<0.01
IBD diagnosis, n (%)				
Crohn's disease	157 (59)	86 (56)	71 (63)	
Ulcerative colitis	69 (26)	46 (30)	23 (20)	0.21
Bechet's disease	41 (15)	22 (14)	19 (17)	
Indication for thiopurine (%)				
Steroid sparing	217 (81)	123 (80)	94 (83)	
Fistula	34 (13)	23 (15)	11 (10)	0.39
Active disease	16 (6)	8 (3)	8 (7)	
Initial AZA dose in mg/kg, mean	1.0	0.97	0.99	0.72
Body weight in kg, mean	56	59	53	<0.01
Baseline WBC in /mm^3, mean	7,326	7,347	7,297	0.83
Baseline neutrophil in /mm^3, mean	5,283	4,880	5,822	0.11
Baseline lymphocyte in /mm^3, mean	1,687	1,753	1,596	0.44
Baseline hemoglobin in g/dl, mean	12.1	12.5	11.6	<0.01
Baseline platelet in $\times 10^3$/mm^3, mean	354	358	349	0.54

AZA, azathioprine; IBD, inflammatory bowel disease; WBC, white blood cell

Table 3. Ten significant SNPs associated with thiopurine-induced leucopenia in the GWAS (n = 267)

SNP	Genomic position (bp)	Nearby gene (distance, bp)	Location	Minor allele	MAF		*P* value[a]	OR[b] (95% CI)
					Leukopenia	No leukopenia		
rs2834826	21q22.12 (36564651)	*RUNX1* (143056), *LOC100506403* (180154)	intergenic	T	56.7%	33.1%	5.8×10^{-8}	2.6 (1.9-3.8)
rs2248159	21q22.12 (36576601)	*RUNX1* (155006), *LOC100506403* (168204)	intergenic	A	59.3%	37.7%	7.5×10^{-7}	2.4 (1.7-3.4)
rs2834832	21q22.12 (36573013)	*RUNX1* (151418), *LOC100506403* (171792)	intergenic	T	61.6%	40.2%	1.1×10^{-6}	2.4 (1.7-3.4)
rs16952570	16q12.2 (53865730)	*FTO*	intronic	C	41.2%	21.8%	1.3×10^{-6}	2.5 (1.7-3.7)
rs1969003	3q23 (31407745)	*GADL1* (471592), *STT3B* (166746)	intergenic	C	50.4%	30.1%	1.8×10^{-6}	2.4 (1.7-3.4)
rs7605946	2p24.1 (23044808)	*LOC645949* (1111284), *KLHL29* (563490)	intergenic	T	57.1%	36.3%	1.9×10^{-6}	2.3 (1.6-3.3)
rs2065296	21q22.12 (36594858)	*RUNX1* (173263), *LOC100506403* (149947)	intergenic	A	50.4%	30.2%	2.1×10^{-6}	2.4 (1.6-3.4)
rs2064056	21q22.12 (36595108)	*RUNX1* (173513), *LOC100506403* (149697)	intergenic	T	56.2%	35.7%	2.5×10^{-6}	2.3 (1.6-3.3)
rs10521308	16q12.2 (53865126)	*FTO*	intronic	A	40.6%	21.8%	2.6×10^{-6}	2.5 (1.7-3.6)
rs10178635	2p24.1 (23040566)	*LOC645949* (1107042), *KLHL29* (567732)	intergenic	T	57.5%	37.3%	3.8×10^{-6}	2.3 (1.6-3.2)

SNP, single nucleotide polymorphism; OR, odds ratio; MAF, minor allele frequency

[a]Results from the chi-square test. [b]Estimated for the risk allele from a 2x2 allele frequency table.

High-Throughput Sequencing for Fine Mapping

Next, we performed high-throughput sequencing in the discovery cohort (n = 331 including 267 patients analyzed in the GWAS; 132 leukopenia and 199 no leukopenia) of the coding regions surrounding the top two markers as well as of genes previously reported as related to thiopurine metabolism and transport (Table 1). On average, the median depth of coverage was 421X, and greater than 97% of the targeted regions were covered by ≥ 20X. From a total of 273 coding variants, we identified 48 nonsynonymous variants with a mean allele frequency (MAF) ≥ 1% (Table 4, Figure 9). Among these coding variants, a missense variant in *FTO* (c.400 G>A; p.Ala134Thr; rs79206939) showed the strongest association with thiopurine-induced leukopenia (P = 0.0013; odds ratio [OR] = 6.3; Table 4). Based on D' analysis, *FTO* p.A134T (rs79206939) was located within a single LD block containing the *FTO* intron 3 variant (rs16952570) and all samples with rs79206939 had rs16952570 as well (D' = 1.0, Figure 6B). However, r^2 was low (r^2 = 0.06) because the allele frequency of rs79206939 (MAF = 0.023) was much smaller than that of rs16952570 (MAF = 0.296). Notably, the genotype frequency of *FTO* p.A134T in patients who experienced thiopurine-induced leukopenia was 9.0%, but its frequency in individuals without leukopenia was only 1.5% (Figure 6C). The median time to leukopenia was 6 months (95% confidence interval [CI], 3.5 to 6.5 months) in individuals with *FTO* p.A134T (Figure 6D). However, individuals with wild-type FTO did not reach the median level during a 36-month observation period, indicating a lower probability of leukopenia compared with individuals with *FTO* p.A134T (log-rank P < 0.001, figure 6D).

FTO p.A134T was only observed in subjects of East Asian ancestry (MAF, 2.2%) in the 1000 Genomes Databases (Table 5). Likewise, *FTO* p.A134T was observed in 5.1% (genotype frequency, 43 of 842) of healthy Korean individuals (Figure 10A). Other *FTO* coding variants occurred at a genotype frequency of less than 1% in Korean individuals and did not appear to be

involved in thiopurine-induced leukopenia. For example, the second most common nonsynonymous variant, *FTO* D144N, was observed in only one control and in none of the leukopenia subjects (Figure 10B). Interestingly, a multivariate analysis using demographic characteristics identified as significant in the univariate analysis (Table 2; female sex, initial body weight, and baseline hemoglobin level) and the genotype data of *FTO* p.A134T and *RUNX1* rs2834826 indicated that *FTO* and *RUNX1* genotypes are more significant determinant for thiopurine-induced leukopenia (Table 6).

Table 4. Association results of the coding variants identified from high-throughput sequencing in the discovery cohort (n = 331)

SNP	Gene	Amino acid change	Genomic position (bp)		Allele (risk allele)	MAF		*P* value[a]	OR[b] (95% CI)
						Leukopenia	No leukopenia		
rs79206939	*FTO*	p.A134T	16q12.2	(53860052)	A/G (A)	0.045	0.008	0.0013	6.3 (1.8-22.4)
rs11568482	*SLC22A8*	p.I182F	11q12.3	(62763264)	A/T (A)	0.080	0.040	0.0309	2.1 (1.0-4.0)
rs594445	*MOCOS*	p.H703N	18q12.2	(33831189)	A/C (A)	0.235	0.307	0.0438	0.7 (0.5-1.0)
rs2234951	*GSTA2*	p.P110S	6p12.2	(52617738)	A/G (A)	0.171	0.121	0.0707	1.5 (1.0-2.3)
rs4148356	*ABCC1*	p.R723Q	16p13.11	(16177275)	A/G (A)	0.080	0.050	0.1257	1.6 (0.9-3.1)
rs683369	*SLC22A1*	p.L160F	6q25.3	(160551204)	G/C (G)	0.159	0.118	0.1300	1.4 (0.9-2.2)
rs4646227	*SLC15A1*	p.G419A	13q32.3	(99358401)	G/C (G)	0.049	0.078	0.1474	0.6 (0.3-1.2)
rs10868138	*SLC28A3*	p.Y113C	9q21.33	(86917301)	C/T (C)	0.098	0.073	0.2422	1.4 (0.8-2.4)
rs2231137	*ABCG2*	p.V12M	4q22.1	(89061114)	T/C (T)	0.250	0.289	0.2710	0.8 (0.6-1.2)
rs143935618	*AOX1*	p.I598N	2q33.1	(201485461)	A/T (A)	0.015	0.028	0.2905	0.5 (0.2-1.7)
rs10883841	*NT5C2*	p.T3A	10q24.33	(104934709)	C/T (C)	0.061	0.043	0.3003	1.4 (0.7-2.9)
rs74837985	*GSTM1*	p.K173N	1p13.3	(110233138)	C/G (C)	0.341	0.379	0.3136	0.8 (0.6-1.2)
rs1057251	*MOCOS*	p.V867A	18q12.2	(33848581)	C/T (C)	0.038	0.025	0.3479	1.5 (0.6-3.7)
rs2180314	*GSTA2*	p.S112T	6p12.2	(52617731)	C/G (C)	0.296	0.264	0.3728	1.2 (0.8-1.7)
rs3731722	*AOX1*	p.H1297R	2q33.1	(201534389)	G/A (G)	0.083	0.065	0.3817	1.3 (0.7-2.3)
rs2297322	*SLC15A1*	p.S117N	13q32.3	(99376181)	T/C (T)	0.409	0.442	0.3992	0.9 (0.6-1.2)
rs2273697	*ABCC2*	p.V417I	10q24.2	(101563815)	A/G (A)	0.091	0.073	0.4022	1.3 (0.7-2.2)
rs2274827	*SLC15A1*	p.R459C	13q32.3	(99356584)	A/G (A)	0.027	0.038	0.4323	0.7 (0.3-1.7)
rs3765534	*ABCC4*	p.E757K	13q32.1	(95815415)	T/C (T)	0.076	0.060	0.4344	1.3 (0.7-2.4)
rs11568658	*ABCC4*	p.G187W	13q32.1	(95863008)	A/C (A)	0.152	0.131	0.4474	1.2 (0.8-1.9)
rs11854484	*SLC28A2*	p.P22L	15q21.1	(45545478)	T/C (T)	0.038	0.050	0.4536	0.7 (0.3-1.6)

SNP	Gene	Amino acid change	Genomic position (bp)		Allelle (risk allele)	MAF		*P* value[a]	OR[b] (95% CI)
						Leukopenia	No leukopenia		
rs199816990	*GSTM1*	p.R96L	1p13.3	(110231874)	T/G (T)	0.015	0.023	0.4981	0.7 (0.2-2.2)
rs45547640	*XDH*	p.N1109T	2p23.1	(31569660)	G/T (G)	0.057	0.045	0.5022	1.3 (0.6-2.6)
rs6577	*GSTA2*	p.E210A	6p12.2	(52615415)	G/T (G)	0.125	0.143	0.5031	0.9 (0.5-1.4)
rs2282143	*SLC22A1*	p.P341L	6q25.3	(160557643)	T/C (T)	0.182	0.163	0.5356	1.1 (0.8-1.7)
rs202002774	*GSTM1*	p.M105T	1p13.3	(110231901)	C/T (C)	0.008	0.013	0.5390	0.6 (0.1-3.1)
rs55754655	*AOX1*	p.N1135S	2q33.1	(201526330)	G/A (G)	0.004	0.008	0.5421	0.5 (0.05-4.8)
rs2032582	*ABCB1*	p.S893A	7q21.12	(87160618)	C/A (C)	0.379	0.357	0.5649	1.1 (0.8-1.5)
rs7023954	*MTAP*	p.V56I	9p21.3	(21816758)	A/G (A)	0.421	0.402	0.6365	1.1 (0.8-1.5)
rs45523133	*XDH*	p.G172R	2p23.1	(31611143)	T/C (T)	0.061	0.053	0.6671	1.2 (0.6-2.3)
rs3744900	*MOCOS*	p.S120N	18q12.2	(33779705)	A/G (A)	0.080	0.088	0.7039	0.9 (0.5-1.6)
rs1127354	*ITPA*	p.P32T	20p13	(3193842)	A/C (A)	0.125	0.116	0.7142	1.1 (0.7-1.8)
rs1801131	*MTHFR*	p.E429A	1p36.22	(11854476)	G/T (G)	0.189	0.199	0.7723	0.9 (0.6-1.4)
rs2274976	*MTHFR*	p.R594Q	1p36.22	(11850927)	T/C (T)	0.102	0.095	0.7736	1.1 (0.6-1.8)
rs11568681	*ABCC4*	p.L18I	13q32.1	(95953517)	T/G (T)	0.015	0.013	0.7782	1.2 (0.3-4.5)
rs1142345	*TPMT*	p.Y240C	6p22.3	(18130918)	C/T (C)	0.015	0.013	0.7782	1.2 (0.3-4.5)
rs45438191	*SLC22A8*	p.V158A	11q12.3	(62763544)	G/A (G)	0.004	0.005	0.8165	0.8 (0.07-8.3)
rs2271437	*SLC28A2*	p.L163W	15q21.1	(45556120)	G/T (G)	0.057	0.053	0.8218	1.1 (0.5-2.1)
rs623053	*MOCOS*	p.T170I	18q12.2	(33779855)	C/T (C)	0.080	0.083	0.8768	1.0 (0.5-1.7)
rs1801133	*MTHFR*	p.A222V	1p36.22	(11856378)	A/G (A)	0.432	0.427	0.9051	1.0 (0.7-1.4)
rs2274407	*ABCC4*	p.K304N	13q32.1	(95859035)	A/C (A)	0.201	0.204	0.9310	1.0 (0.7-1.4)
rs316019	*SLC22A2*	p.S270A	6q25.3	(160670282)	A/C (A)	0.106	0.108	0.9358	1.0 (0.6-1.6)
rs56350726	*SLC28A3*	p.Y513F	9q21.33	(86900369)	A/T (A)	0.095	0.093	0.9403	1.0 (0.6-1.7)
rs628031	*SLC22A1*	p.M408V	6q25.3	(160560845)	A/G (A)	0.246	0.244	0.9417	1.0 (0.7-1.5)
rs2231142	*ABCG2*	p.Q141K	4q22.1	(89052323)	T/G (T)	0.284	0.286	0.9479	1.0 (0.7-1.4)

SNP	Gene	Amino acid change	Genomic position (bp)	Allelle (risk allele)	MAF		*P* value[a]	OR[b] (95% CI)
					Leukopenia	No leukopenia		
rs1060896	*SLC28A2*	p.S75R	15q21.1 (45554267)	A/C (A)	0.049	0.050	0.9534	1.0 (0.5-2.0)
rs678560	*MOCOS*	p.V358M	18q12.2 (33785093)	G/A (G)	0.087	0.088	0.9709	1.0 (0.6-1.7)
rs2274828	*SLC15A1*	p.V450I	13q32.3 (99356611)	T/C (T)	0.076	0.075	0.9855	1.0 (0.6-1.8)

SNP, single nucleotide polymorphism; OR, odds ratio; MAF, minor allele frequency

[a]Results from the chi-square test. [b]Estimated for the risk allele from a 2x2 allele frequency table.

Table 5. Minor allele frequency of *FTO* p.A134T in 1000 Genome Project, and protein function prediction

Protein change	Genomic change	Minor Allele Frequency (MAF, %) in 1000 Genomes Project					Protein function prediction[a]	
		East Asian	South Asian	African	American	European	LRT	MutationTaster
FTO p.A134T	chr16:53860052 G>A	2.2%	0	0	0	0	Disease causing	Damaging

[a]Functional prediction of amino acid changes was performed using dbNSFP version 2.6 incorporated in the ANNOVAR software with default options. dbNSFP (http://sites.google.com/site/jpopgen/dbNSFP) is a database developed for functional prediction of non-synonymous single-nucleotide variants, which compiles prediction and conservation scores including LRT and Mutation Taster.

Table 6. Multivariate analysis of leukopenia in patients with inflammatory bowel disease

Variable	Category	Univariate analysis			Multivariate analysis		
		OR	95% CI	*P*	OR	95% CI	*P*
Sex	female *v* male (ref)	2.25	1.34 to 3.80	0.002*	1.65	0.91 to 3.01	0.099
Body weight (kg)	<50 *v* ≥50 (ref)	2.17	1.29 to 3.66	0.003*	1.52	0.84 to 2.78	0.168
Baseline hemoglobin (g/dl), mean	<12 *v* ≥12 (ref)	1.85	1.13 to 3.03	0.014*	1.39	0.81 to 2.42	0.234
FTO genotype	p.A134T *v* wild type (ref)	5.98	1.65 to 21.73	0.006*	5.47	1.46 to 20.5	0.012*
RUNX1 genotype	rs2834826 *v* wild type (ref)	2.59	1.51 to 4.42	<0.001	2.42	1.38 to 4.23	0.002*

OR, odds ratio; CI, confidential interval; *$P < 0.05$.

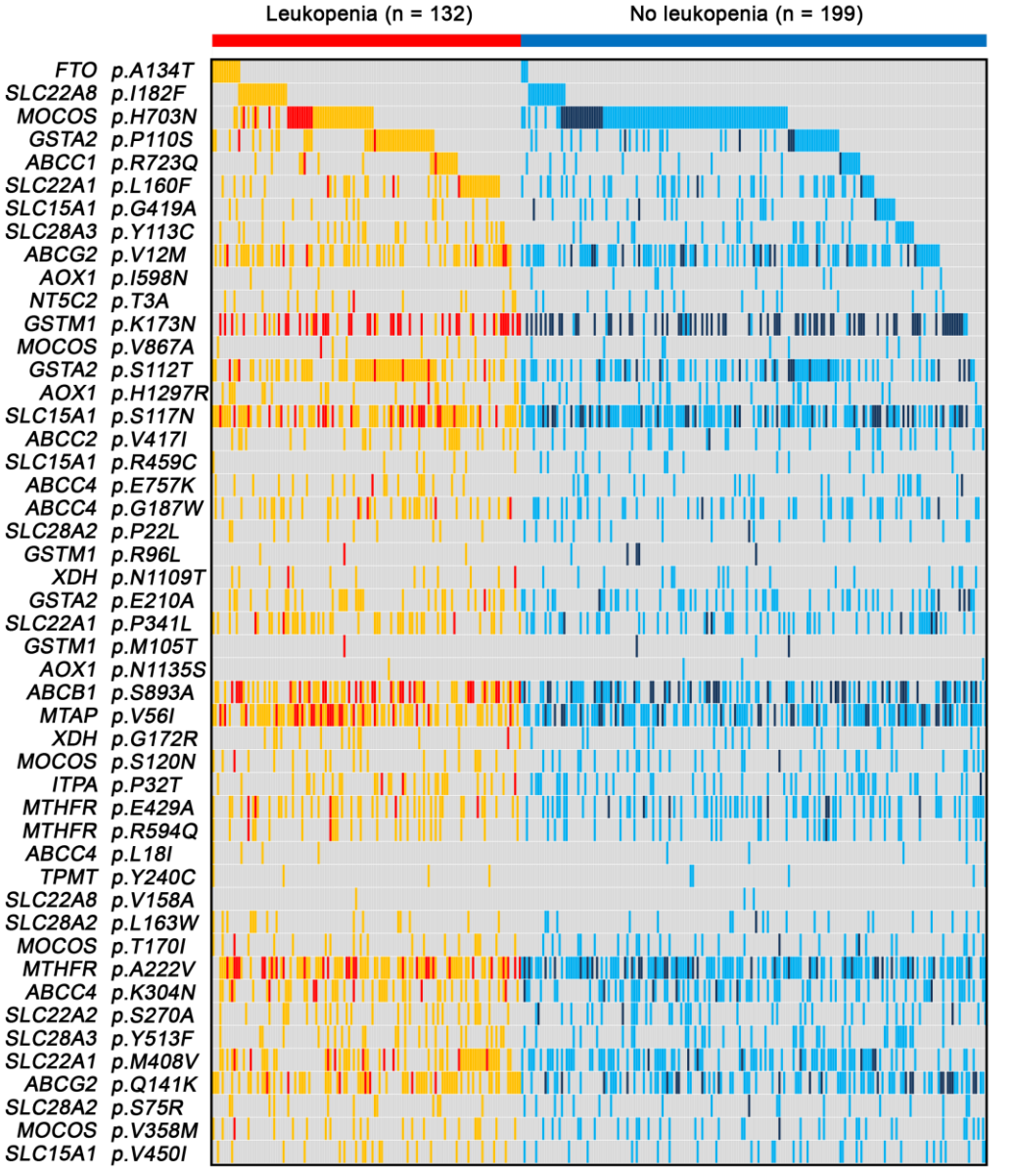


Figure 9. Genotypes of coding variants identified from high-throughput sequencing in the discovery cohort (n = 331). Samples are arranged from left to right. Genotypes of coding variants are annotated according to the color panel below the image.

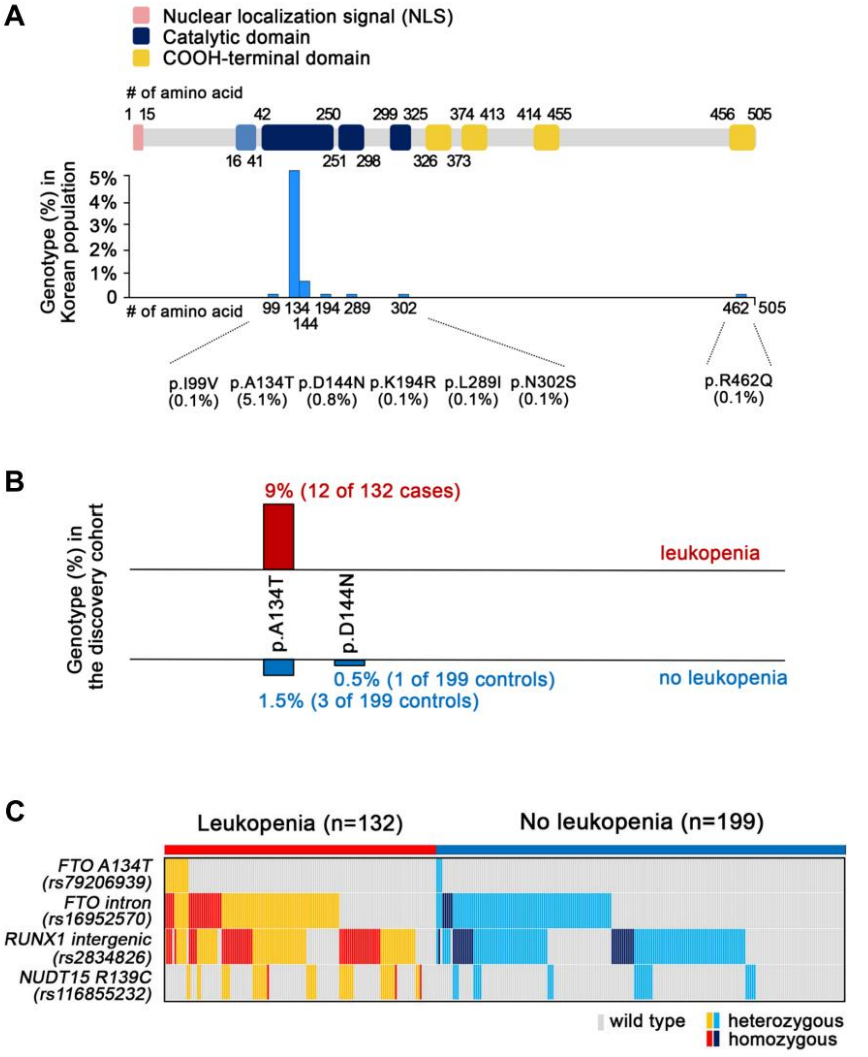


Figure 10. *FTO* coding variants in the Korean population and genotypes associated with thiopurine-induced leukopenia. (A) Frequency of *FTO* coding variants in 842 healthy Korean subjects. FTO is composed of 505 amino acids. (B) Genotype frequencies of patients with (red) and without (blue) leukopenia in the discovery cohort (n = 331). (C) Genotype frequencies of four loci significantly associated with thiopurine-induced leukopenia in the discovery cohort (n = 331; including 267 patients analyzed in the GWAS). Samples are arranged from left to right. Genotypes are annotated according to the color panel below the image.

Replication Study

We next examined whether the association of rs79206939 (*FTO* p.A134T) with thiopurine toxicity is replicated in two additional cohorts. In this replication study, rs116855232 (*NUDT15* p.R139C) recently reported as a causal variant associated with thiopurine-induced leucopenia[15] was included as a reference (table 7). The incidence of leukopenia was 28.4% in the external validation cohort (replication cohort 1; 124 leukopenia and 313 no leukopenia) and 36.7% in the internal validation cohort (replication cohort 2; 121 leukopenia and 209 no leukopenia).

Data from the replication and combined analyses confirmed the novel association of rs79206939 with thiopurine-induced leukopenia identified in the discovery cohort of this study, in addition to that of rs116855232 reported previously (Table 7). The genotype frequency of *FTO* p.A134T was 9-12% in patients with leukopenia including one patient with homozygous variant but was less than 3% in patients without leukopenia in the discovery and replication cohorts (Table 8). Accordingly, leucopenia was observed in 66.7% of samples with *FTO* p.A134T genotype (positive predictive value) and in 32.4% of samples with *FTO* wild-type (Table 9). When analyzed according to IBD subtypes, leucopenia was more frequent in patients with *FTO* p.A134T irrespective of disease subtypes (Table 10). Interestingly, FTO p.A134T was associated not only with thiopurine-induced leukopenia but also other hematologic traits, such as anemia and thrombocytopenia (Figure 11). In addition to *FTO* p.A134T and *NUDT15* p.R139C, the significance of rs16952570 (FTO intron 3) and rs2834826 (*RUNX1* intergenic), which showed an association with thiopurine-induced leukopenia in the discovery cohort, were also reproduced in replication cohorts (Table 11). Results of conditional logistic regression analysis suggest that markers in *FTO*, *NUDT15*, and *RUNX1* are independently associated with thiopurine-induced leukopenia in patients with IBD (Table 12). Interestingly, when *FTO* p.A134T and *NUDT15* p.R139C were combined, leukopenia risk was significantly increased (Table 13).

Table 7. Association of *FTO* p.A134T and *NUDT15* p.R139C with thiopurine toxicity in the discovery and replication cohorts

SNP ID	Genes in or near associated region	Genomic position	Allele (risk allele)	Study	RAF		*P* value[a]	OR (95% CI)[b]	P_{het}[c]
					Leukopenia	No leukopenia			
rs79206939	*FTO p.*A134T	16q12.2	A/G (A)	Discovery	0.045	0.008	1.3×10^{-3}	6.3 (1.8-22.4)	
				Replication1	0.053	0.019	7.5×10^{-3}	2.8 (1.3-6.3)	
				Replication2	0.076	0.015	7.1×10^{-5}	5.5 (2.2-14.1)	
				Combined[d]			**1.3×10^{-8}**	**4.3 (2.5-7.4)**	0.44
rs116855232	*NUDT15* p.R139C	13q14.2	C/T (C)	Discovery	0.159	0.060	3.3×10^{-5}	3.0 (1.7-5.0)	
				Replication1	0.221	0.058	1.0×10^{-12}	4.7 (2.9-7.3)	
				Replication2	0.186	0.050	2.4×10^{-8}	4.3 (2.5-7.4)	
				Combined[d]			**1.4×10^{-22}**	**3.9 (2.9-5.3)**	0.41

Association of *FTO* and *NUDT15* genotypes with thiopurine toxicity was analyzed in the discovery cohort (n = 331, including 267 patients analyzed in the GWAS), replication cohort 1 (n = 437), and replication cohort 2 (n = 330). SNP, single-nucleotide polymorphism; RAF, risk allele frequency; OR, odds ratio; CI, confidence interval; GWAS, genome-wide association study. [a]Results from the chi-square test. [b]Estimated for the risk allele from a 2x2 allele frequency table. [c]Results from the Breslow-Day test. [d]Calculated by the Cochran-Mantel-Haenszel test.

Table 8. Frequency of leucopenia according to *FTO* p.A134T genotype in the discovery cohort (n = 331), replication cohort 1 (n = 437), and replication cohort 2 (n = 330)

	FTO genotype			*P* value
	Homozygote (AA)	Heterozygote (GA)	Wild type (GG)	
No. of leucopenia case (n) in discovery cohort (n = 331)	0/0	12/15 (80%)	120/316 (38%)	0.002
No. of leucopenia case (n) in replication cohort 1 (n = 437)	0/0	13/25 (52%)	111/412 (27%)	0.013
No. of leucopenia case (n) in replication cohort 2 (n = 330)	1/1 (100%)	16/22 (73%)	104/307 (34%)	0.002
All (n = 1,098)	1/1 (100%)	41/62 (66%)	335/1035 (32%)	<0.001

Table 9. Calculation of sensitivity, specificity, positive predictive value, and negative predictive value related to *FTO* p.A134T genotype and leucopenia in patients with inflammatory bowel disease

	Leucopenia	No leucopenia	Total
FTO p.A134T	42	21	63
FTO wild-type	335	700	1035
Total	377	721	1098

Positive-predictive value: 42/63 = 66.7%

Negative-predictive value: 700/1035 = 67.6%

Sensitivity: 42/377 = 11.1%

Specificity: 700/721 = 97.1%

Table 10. Frequency of leucopenia according to *FTO* p.A134T genotype in the Behcet's disease (n = 64), Crohn's disease (n = 719), and Ulcerative colitis (n = 315)

	FTO genotype		
	FTO p.A134T	Wild type (GG)	*P* value
No. of leucopenia case (n) in Behcet's disease	6/6 (100%)	20/58 (34%)	0.003
No. of leucopenia case (n) in Crohn's disease	31/45 (69%)	244/674 (36%)	<0.001
No. of leucopenia case (n) in Ulcerative colitis	5/12 (42%)	71/303 (23%)	0.169
All (n = 1,098)	42/63 (67%)	335/1035 (32%)	<0.001

Table 11. Association of rs16952570 (*FTO* intron 3) and rs2834826 (*RUNX1* intergenic) with thiopurine toxicity in the discovery and replication cohorts (n = 846)

SNP ID	Genes in or near associated region	Genomic position	Allele (risk allele)	Study	RAF		P valuea	OR (95% CI)b	Phetc
					Case	Control			
rs16952570	FTO intron 3	16q12.2	C/T (C)	Discovery	0.402	0.226	1.3 × 10-6	2.3 (1.6-3.2)	
				Replication1	0.373	0.271	4.1 × 10-2	1.6 (1.0-2.5)	
				Replication2	0.403	0.287	2.4 × 10-3	1.7 (1.2-2.4)	
				Combinedd			4.2 × 10-9	1.9 (1.5-2.3)	0.33
rs2834826	RUNX1 intergenic	21q22.12	T/C (T)	Discovery	0.549	0.349	3.6 × 10-7	2.3 (1.7-3.1)	
				Replication1	0.582	0.424	3.3 × 10-3	1.9 (1.2-2.9)	
				Replication2	0.492	0.404	3.0 × 10-2	1.4 (1.0-2.0)	
				Combinedd			3.3 × 10-9	1.8 (1.5-2.2)	0.12

Association of FTO and RUNX1 genotypes with thiopurine toxicity was analyzed in the discovery cohort (n = 331, including 267 patients analyzed in the GWAS), replication cohort 1 (n = 185), and replication cohort 2 (n = 330). SNP, single-nucleotide polymorphism; RAF, risk allele frequency; OR, odds ratio; CI, confidence interval; GWAS, genome-wide association study. aResults from the chi-square test. bEstimated for the risk allele from a 2x2 allele frequency table. cResults from the Breslow-Day test. dCalculated by the Cochran-Mantel-Haenszel test.

Table 12. Conditional logistic regression analysis of *FTO*, *NUDT15*, and *RUNX1* SNPs

Test SNP	Conditioned SNP	*P* value	OR (95% CI)
rs79206939 (*FTO* p.A134T)	rs116855232 (*NUDT15* p.R139C)	1.1×10^{-7}	6.0 (3.1-11.6)
rs116855232 (*NUDT15* p.R139C)	rs79206939 (*FTO* p.A134T)	6.0×10^{-16}	4.5 (3.1-6.5)
rs79206939 (*FTO* p.A134T)	rs2834826 (*RUNX1* intergenic)	2.5×10^{-6}	4.8 (2.5-9.2)
rs2834826 (*RUNX1* intergenic)	rs79206939 (*FTO* p.A134T)	9.6×10^{-8}	1.7 (1.4-2.1)
rs2834826 (*RUNX1* intergenic)	rs116855232 (*NUDT15* p.R139C)	4.0×10^{-8}	1.8 (1.5-2.2)
rs116855232 (*NUDT15* p.R139C)	rs2834826 (*RUNX1* intergenic)	9.3×10^{-16}	4.6 (3.1-6.6)

Data shown in table 7 (*FTO* and *NUDT15*; n = 1,035) and table 11 (*FTO* and *RUNX1*, *RUNX1* and *NUDT15*; n = 846) are used. OR, odds ratio; CI, confidence interval.

Table 13. Combined effects of risk alleles of *FTO*, *NUDT15*, and *RUNX1* SNPs on thiopurine-induced leukopenia

SNP	MAF (leukopenia)	MAF (no leukopenia)	*P* value	OR (95% CI)
rs79206939 (*FTO* p.A134T)	0.058	0.015	1.81×10^{-8}	4.1 (2.4-7.0)
rs116855232 (*NUDT15* p.R139C)	0.188	0.056	3.20×10^{-22}	3.9 (2.9-5.2)
rs79206939 and rs116855232	0.015	0.002	4.71×10^{-7}	7.1 (2.0-25.5)
rs79206939 (*FTO* p.A134T)	0.062	0.013	1.90×10^{-8}	5.2 (2.8-9.8)
rs2834826 (*RUNX1* intergenic)	0.534	0.388	4.32×10^{-9}	1.8 (1.5-2.2)
rs79206939 and rs2834826	0.050	0.009	7.59×10^{-8}	6.1 (2.9-12.9)

Data shown in table 4 (*FTO* and *NUDT15*; n = 1,098) and supplementary table S7 (*FTO* and *RUNX1*; n = 846) are used. OR, odds ratio; CI, confidence interval.

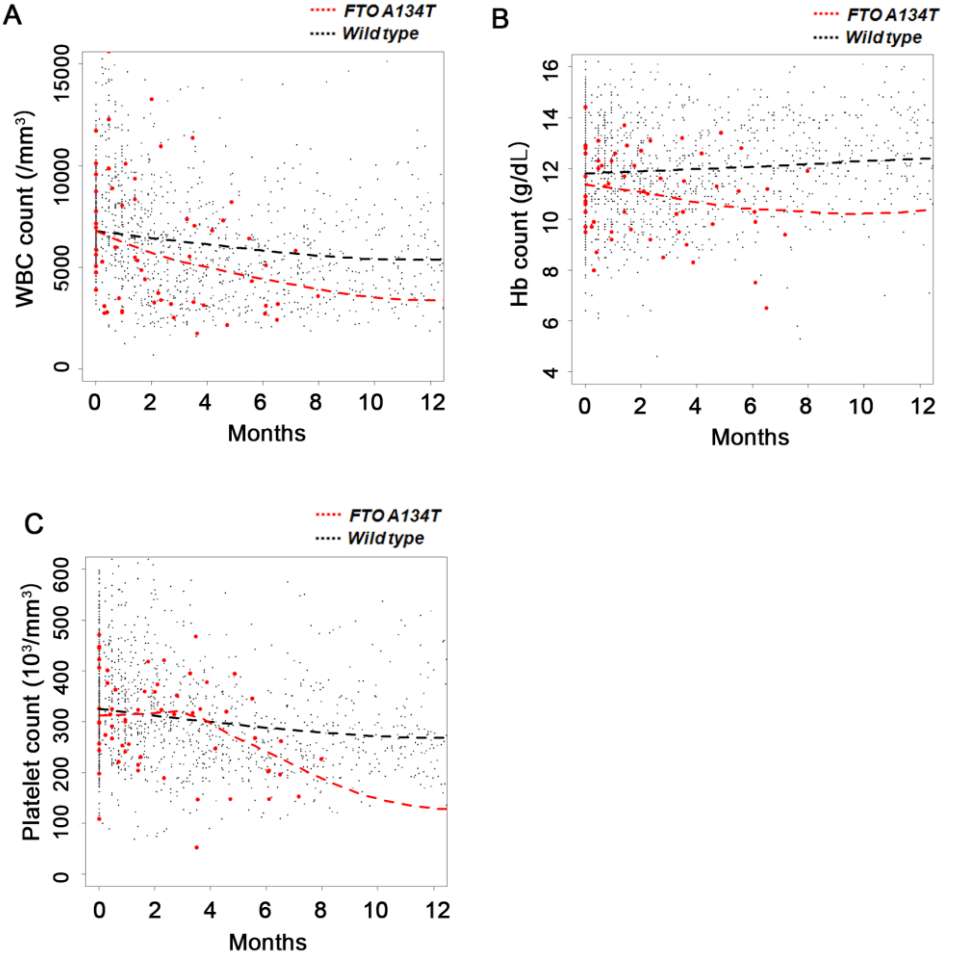


Figure 11. Hematological traits of *FTO* A134T and *FTO* wild type during thiopurine therapy. Hematologic data (*y*-axis) are plotted according to time (*x*-axis). Each dot (*FTO* A134T in *red* and *FTO* wild type in *black*) represents hematologic data at the indicated time. For each genotype, the dotted line was plotted using a Loess (locally weighted regression) fit. (*A*) white blood cell count; (*B*) hemoglobin concentration; (*C*) platelet count.

Kinetic Analyses of FTO p.Ala134Thr Demethylase Activity

FTO is a nuclear protein encoding a 2-oxyglutarate and Fe(II)-dependent demethylase and is closely related to the mammalian AlkB homologues ABH1 and ABH2.[30] The crystal structure of FTO shows an N-terminal catalytic domain and a C-terminal domain of unknown function, and FTO A134 is located within an alpha helix in the N-terminal catalytic domain (Figure 12A). Interestingly, FTO A134 is highly conserved among mammals (Figure 12B). In addition, protein prediction algorithms suggest that the p.A134T mutation may detrimentally impact FTO protein function (Table 5). These findings highly suggest that the p.A134T variation may affect FTO structure or function.

We first examined whether the p.A134T variation affects the cellular expression of FTO. However, FTO expression in Jurkat human lymphocytes was not altered by the p.A134T mutation (Figure 12C). We then evaluated the demethylase activity of purified recombinant FTO proteins to investigate the functional consequences of the p.A134T variation using FTO wild type and FTO p.R316Q (rs121918214) as positive and negative controls, respectively. It was shown previously that the FTO p.R316Q inactivating mutation causes a severe developmental polymalformative syndrome in humans.[31] Interestingly, the demethylase activity of the FTO p.A134T mutant protein was 65% lower than that of wild-type FTO (Figure 12D). As expected, the FTO p.R316Q mutant protein exhibited no detectable demethylase activity.

Next, because the heterozygous *FTO* variant encoding p.A134T was associated with thiopurine-induced leukopenia (Figure 10C), we examined whether FTO p.A134T inhibits the function of FTO wild type. The demethylase activity of FTO was measured in the presence of FTO wild type with or without the same amount of FTO p.A134T. As shown in Figure 12E, FTO p.A134T had no effect on the demethylase activity of FTO wild type. Taken together, the results of the *in vitro* studies indicate that p.A134T is a hypomorphic variation in FTO and does not exert a dominant-negative effect on FTO wild type.

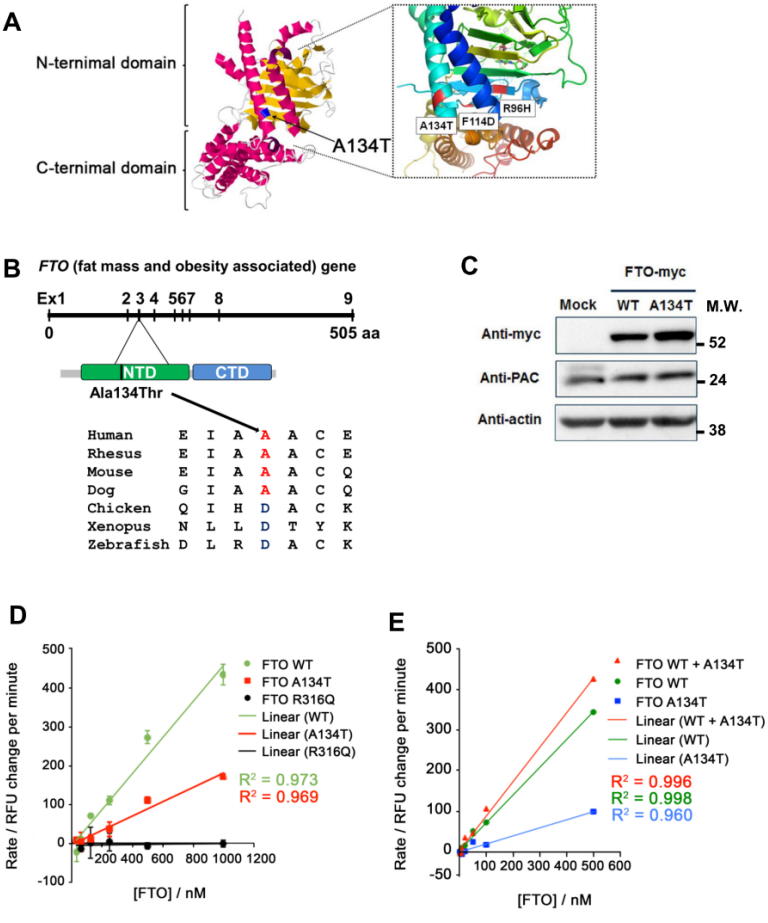


Figure 12. Kinetic analyses evaluating FTO p.A134T demethylase activity. (A) Stereo view of the overall structure of FTO. The loss-of-function mutation FTO R96H within the substrate recognition lid, F114D involved in interactions between N- and C-terminal domains, and p.A134T are shown (*red*). (B) Schematic diagram of the domain structure of FTO and amino acid at position 134 among vertebrates. (C) Immunoblot of FTO wild type and FTO p.A134T proteins. Jurkat lymphocytes were transfected with lentiviral vectors encoding myc-tagged FTOs and selected with puromycin. Transferred proteins were blotted with anti-myc, anti-PAC (puromycin N-acetyltransferase), and anti-actin (β-actin) antibodies. (D) Demethylase activity was measured in FTO wild type (*green*), p.A134T (*red*), and p.R316Q (*black*; mutant lacking demethylase activity) using the FTO 3-methyluracil demethylation assay. (*E*) FTO 3-methyluracil demethylation assay performed with FTO wild type alone (*green*), FTO p.A134T alone (*blue*), and both FTO wild type and FTO p.A134T (*red*) to evaluate whether FTO p.A134T inhibits the function of FTO wild type.

***In Vivo* Thiopurine Treatment in *Fto*$^{-/-}$ Mice**

Using *Fto*-deficient mice, we next investigated whether reduced FTO function increases the risk of thiopurine-induced myelosuppression. Mice received AZA or 6-MP for 2 weeks, after which peripheral blood CBCs and bone marrow examinations were conducted (n = 6 per group). Notably, administration of thiopurines evoked a more severe leukopenia in *Fto*-deficient mice. For example, in mice treated with AZA (30 mg/kg, once daily p.o.) for 2 weeks, the number of peripheral blood WBCs decreased to 59% of the baseline number in wild-type (*Fto*$^{+/+}$) mice. The decrease in WBC count was more profound in both *Fto*$^{+/-}$ (42% of baseline, $P < 0.05$) and *Fto*$^{-/-}$ mice (31% of baseline, $P < 0.01$) (Figure 13A). Similar results were observed in experiments with 6-MP. Treatments with 6-MP (10 mg/kg, once daily i.v.) for 2 weeks reduced the WBC count to 67% of the baseline count in *Fto*$^{+/+}$ mice, and this value was reduced further to 53% in *Fto*$^{+/-}$ mice ($P < 0.01$) and 40% in *Fto*$^{-/-}$ mice ($P < 0.01$) (Figure 13B).

Consistent with the effect on peripheral WBC counts, *Fto* deficiency induced a more profound decrease in bone marrow cellularity with thiopurine treatment (Figure 13C-13F). Treatment with AZA evoked a 16% reduction in bone marrow cellularity in *Fto*$^{+/+}$ mice, a 38% reduction in *Fto*$^{+/-}$ mice ($P < 0.01$), and 50% reduction in *Fto*$^{-/-}$ mice ($P < 0.01$) (Figure 13C and 13E). Treatment with 6-MP did not evoke discernible reduction in bone marrow cellularity in *Fto*$^{+/+}$ mice. However, the same treatment reduced the cellularity by 19% ($P < 0.05$) and 31% ($P< 0.01$) in *Fto*$^{+/-}$ and *Fto*$^{-/-}$ mice, respectively (Figure 13D and 13F). The results of *in vivo* experiments in mice strongly suggest that *Fto* deficiency increases susceptibility to myelosuppression.

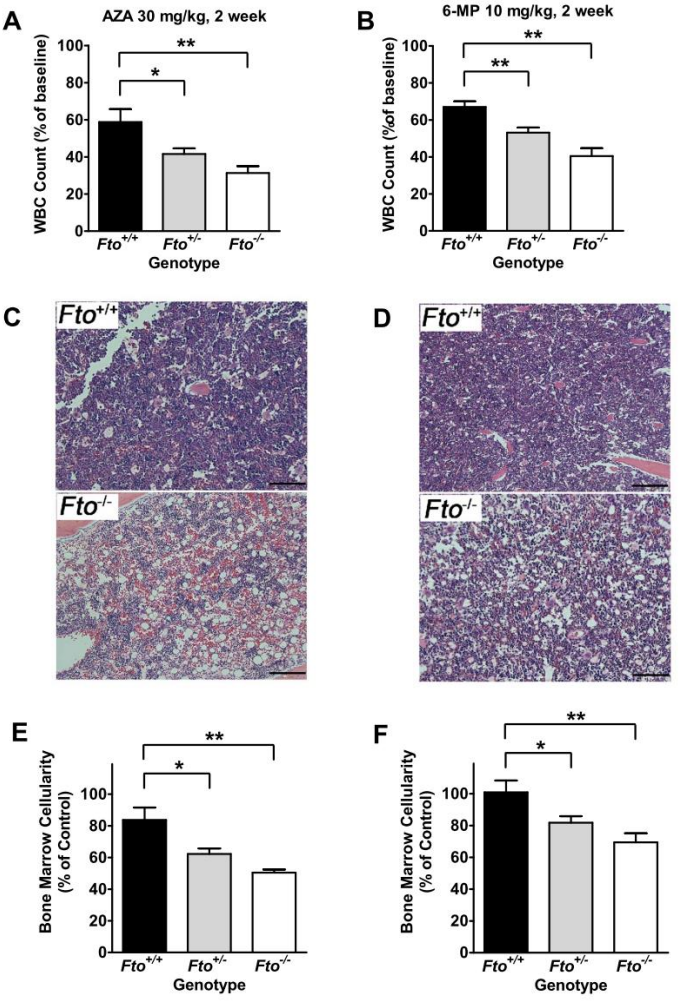


Figure 13. *In vivo* thiopurine treatment in *Fto*-deficient mice. (A and B) Effect of *Fto* genotype on thiopurine-induced leukopenia was examined using peripheral blood samples. Mice received AZA (*A*; 30 mg/kg, p.o.) or 6-MP (*B*; 10 mg/kg, i.p.) daily for 2 weeks. Blood samples were collected before and after thiopurine treatment. The white blood cell (WBC) count as a percentage of baseline at 2 weeks is shown. Thiopurine-induced leukopenia was more severe in $Fto^{+/-}$ and $Fto^{-/-}$ mice. (C–F) Effect of *Fto* genotype on thiopurine-induced bone marrow suppression. Mice received AZA (C and E; 30 mg/kg, p.o.) or 6-MP (D and F; 10 mg/kg, i.p.) daily for 2 weeks, after which bone marrow cellularity was examined. Representative images of femur sections stained with hematoxylin-eosin are presented in C and D (scale bars, 200 μM). Results of multiple experiments are summarized in E and F. Thiopurine-induced bone marrow suppression was more pronounced in $Fto^{+/-}$ and $Fto^{-/-}$ mice. Error bars indicate mean ± SEM (n=6). *, $P < 0.05$; **, $P < 0.01$.

IV. DISCUSSION

Predicting toxicity in individuals with an at-risk genotype is crucial in thiopurine therapy due to the narrow therapeutic index.[1,9] Integrating genotype in risk classification can prevent life-threatening toxicity and enhance therapeutic efficacy. In the present study, we performed a GWAS to identify novel variants associated with thiopurine-induced leukopenia. In a subsequent high-throughput sequencing analysis, we identified a novel coding variant, *FTO* p.A134T, located within the same LD block containing the *FTO* intronic variant associated with thiopurine-induced leukopenia in the GWAS. Finally, the functional consequences of the FTO p.A134T mutation were examined both *in vitro* and *in vivo*. The integrated results strongly suggest that the hypomorphic *FTO* p.A134T variant confers susceptibility to thiopurine-induced leukopenia in patients with IBD.

The present findings provide new insights into the physiological role of FTO. In this study, most individuals with the *FTO* p.A134T allele were $FTO^{A134T/wild}$ heterozygotes (Table 8). Interestingly, heterozygote loss of FTO was sufficient to increase the risk of thiopurine-induced leukopenia in association studies (OR = 4.3; Figure 10C, table 7). This observation was comparable to results of mouse *in vivo* experiments in which AZA and 6-MP evoked more severe leukopenia and bone marrow suppression in $Fto^{+/-}$ mice compared with wild-type mice (Figure 13). Because FTO p.A134T did not exhibit a dominant negative effect on wild-type FTO (Figure 12E), a partial reduction in FTO activity is anticipated in individuals with $FTO^{A134T/wild}$. A partial reduction in protein function can increase disease risk when the protein plays a critical role in the disease pathway. For example, the p.R254W variation in chymotrypsin C reduces the enzyme activity by approximately 50% and significantly increases the risk of chronic pancreatitis in heterozygous individuals (OR = 4.6).[32] Therefore, FTO function appears to be critical for thiopurine-induced damage repair in hematopoietic cells.

FTO has a nucleotide demethylase activity, although the physiological significance of this activity is uncertain. One of the remaining questions in this study is how FTO p.A134T contributes to thiopurine-induced leukopenia. Considering thiopurine metabolism and the nucleotide modifying activity of FTO, several possibilities exist as underlying mechanisms of FTO p.A134T–associated thiopurine-induced toxicity. For example, the demethylase activity of FTO can abrogate methyl-TIMP, a thiopurine metabolite that acts as a potent purine synthesis inhibitor.[1] FTO may also participate in the repair of DNA and RNA damages induced by thiopurines. FTO belongs to the AlkB family of Fe(II)/α-ketoglutarate–dependent dioxygenases that reverse DNA alkylation damage.[33] Demethylation by FTO can repair 3-methylthymine in single-stranded DNA and 3-methyluracil or N^6-methyladenosine (m^6A) in RNA.[33] Lastly, FTO may also participate in hematopoiesis or protect cells from thiopurine toxicity by regulating the expression of other genes. A recent report highlighted the role of FTO in RNA modification.[34] As "erasers" that demethylate m^6A in RNA, FTO has been suggested to regulate global gene expression post-transcriptionally.[35] Interestingly, in a gene expression analysis in mice overexpressing *Fto*, *Tpmt* was shown to be upregulated.[36] It would be intriguing to identify the precise molecular target of FTO in thiopurine toxicity, as this might elucidate the physiological significance of FTO's demethylase activity. It is well known that some markers in *FTO* regions are associated with obesity. For example, genetic association studies in European cohorts showed that rs9939609 and rs1421085 in *FTO* intron 1 are strongly associated with obesity.[37,38] However, *FTO* A134T in exon 3 did not show linkage disequilibrium with rs9939609 and rs1421085, suggesting that *FTO* A134T is not associated with obesity.

In the GWAS, an intergenic region proximal to *RUNX1* was associated with thiopurine-induced leukopenia (Figure 6, table 3). Although the effect size was relatively small (OR = 1.8), rs2834826 in this region was consistently

significant in the discovery and replication cohorts (P=3.3×10^{-9}, table 11). Considering the fact that the RUNX1 transcription factor plays an essential role in hematopoiesis,[39] identification of a functional RUNX1 variant linked to rs2834826 would be clinically meaningful. However, in subsequent high-throughput sequencing, we were unable to find the causative coding variant in the *RUNX1* exome area. A rare variant of *RUNX1*, p.L445P, appeared to be enriched in patients with leukopenia (P = 0.04) but was not located in the same LD block containing rs2834826. Further research will be required to properly interpret these noncoding variants associated with thiopurine-induced leukopenia.

The value of *TPMT* genotyping is limited in East Asians. In our study, the **2*, **3A*, and **3B TPMT*-null polymorphisms were not detected in Korean patients with IBD, whereas in individuals of European descent with thiopurine-induced leukopenia, up to 25% have these variations.[15-17] Meanwhile, 3.8% of Koreans have the *TPMT*3C* polymorphism (*TPMT* p.Y240C) that results in moderate enzyme activity. However, the frequency of *TPMT*3C* was not different between those with and without leukopenia (P = 0.778; table 3), suggesting that *TPMT*3C* is not an important determinant. Previous studies have suggested several non-*TPMT* genetic determinants of thiopurine toxicity, including variations in *ITPA*, *ABCC4*, *ABCB1/MDR1*, *MTHFR*, and *HPRT1*.[13,40,41] However, we did not identify any significant polymorphisms in these genes, which showed frequency differences between the leukopenia samples and no leukopenia samples.

In the initial GWAS of present study, none of the markers reached the conservative genome-wide significance of P<5×10^{-8} in the discovery cohort possibly due to the small number of IBD cases studies. However, some LD blocks showed potential association with thiopurine toxicity, especially those in 16q12 and 21q22 regions (Figure 6a). Notably, subsequent targeted NGS analysis identified *FTO* p.A134T as a causative SNP for 16q12 (Table 3) and

validation studies successfully replicated the association between *FTO* p.A134T and thiopurine toxicity (P=1.3×10^{-8}, table 7). These results have confirmed the utility of using small numbers of well characterized cases to identify clinically useful genetic risk factors for adverse responses to drug therapy by GWAS methodologies.

V. CONCLUSION

We provide evidence that *FTO* p.A134T is associated with thiopurine-induced leukopenia in patients with IBD. Inter-individual variation in susceptibility to thiopurine poses a significant therapeutic obstacle for clinicians. Although further studies including prospective trials are required to validate the clinical utility of the *FTO* variant, our findings of identifying the risk allele may be incorporated into the genotype-guided thiopurine dosing recommendations in the precision medicine era.

REFERENCES

1. Karran P, Attard N. Thiopurines in current medical practice: molecular m echanisms and contributions to therapy-related cancer. Nat Rev Cancer 2 008;8:24-36.
2. Chande N, Tsoulis DJ, MacDonald JK. Azathioprine or 6-mercaptopurine for induction of remission in Crohn's disease. Cochrane Database Syst R ev 2013;4:CD000545.
3. Timmer A, McDonald JW, Tsoulis DJ, Macdonald JK. Azathioprine and 6-mercaptopurine for maintenance of remission in ulcerative colitis. Coch rane Database Syst Rev 2012;9:CD000478.
4. Fraser AG, Orchard TR, Jewell DP. The efficacy of azathioprine for thet reatment of inflammatory bowel disease: a 30 year review. Gut 2002;50: 485-9.
5. Kirschner BS. Safety of azathioprine and 6-mercaptopurine in pediatric p atients with inflammatory bowel disease. Gastroenterology 1998;115:813-2 1.
6. Connell WR, Kamm MA, Ritchie JK, Lennard-Jones JE. Bone marrow t oxicity caused by azathioprine in inflammatory bowel disease: 27 years of experience. Gut 1993;34:1081-5.
7. Kim JH, Cheon JH, Kim WH. [The frequency and the course of the ad verse effects of azathioprine/6-mercaptopurine treatment in patients with i nflammatory bowel disease]. Korean J Gastroenterol 2008;51:291-7.
8. Kim JH, Cheon JH, Hong SS, Eun CS, Byeon JS, Hong SY, et al. Infl uences of thiopurine methyltransferase genotype and activity on thiopurin e-induced leukopenia in Korean patients with inflammatory bowel diseas e: a retrospective cohort study. J Clin Gastroenterol 2010;44:e242-8.
9. Gisbert JP, Gomollon F. Thiopurine-induced myelotoxicity in patients wit h inflammatory bowel disease: a review. Am J Gastroenterol 2008;103:17 83-800.
10. Gearry RB, Barclay ML, Burt MJ, Collett JA, Chapman BA. Thiopurine drug adverse effects in a population of New Zealand patients with infla mmatory bowel disease. Pharmacoepidemiol Drug Saf 2004;13:563-7.
11. Warman JI, Korelitz BI, Fleisher MR, Janardhanam R. Cumulative experi ence with short- and long-term toxicity to 6-mercaptopurine in the treatm ent of Crohn's disease and ulcerative colitis. J Clin Gastroenterol 2003;3 7:220-5.
12. Lichtenstein GR, Abreu MT, Cohen R, Tremaine W, American Gastroent erological A. American Gastroenterological Association Institute technicalr eview on corticosteroids, immunomodulators, and infliximab in inflammat ory bowel disease. Gastroenterology 2006;130:940-87.
13. Zaza G, Cheok M, Krynetskaia N, Thorn C, Stocco G, Hebert JM, et al Thiopurine pathway. Pharmacogenet Genomics 2010;20:573-4.
14. Jung YS, Cheon JH, Park JJ, Moon CM, Kim ES, Lee JH, et al. Correl ation of genotypes for thiopurine methyltransferase and inosine triphospha te pyrophosphatase with long-term clinical outcomes in Korean patients with inflammatory bowel diseases during treatment with thiopurine drugs. J Hum Genet 2010;55:121-3.

15. Yang SK, Hong M, Baek J, Choi H, Zhao W, Jung Y, et al. A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. Nat Genet 2014;46:1017-20.
16. Colombel JF, Ferrari N, Debuysere H, Marteau P, Gendre JP, Bonaz B, et al. Genotypic analysis of thiopurine S-methyltransferase in patients wit h Crohn's disease and severe myelosuppression during azathioprine therap y. Gastroenterology 2000;118:1025-30.
17. Dewit O, Moreels T, Baert F, Peeters H, Reenaers C, de Vos M, et al. Limitations of extensive TPMT genotyping in the management of azathio prine-induced myelosuppression in IBD patients. Clin Biochem 2011;44:1 062-6.
18. Kakuta Y, Naito T, Onodera M, Kuroha M, Kimura T, Shiga H, et al. NUDT15 R139C causes thiopurine-induced early severe hair loss and leu kopenia in Japanese patients with IBD. Pharmacogenomics J 2015; doi:1 0.1038/tpj.2015.43.
19. Grunnet LG, Nilsson E, Ling C, Hansen T, Pedersen O, Groop L, et al. Regulation and function of FTO mRNA expression in human skeletal mu scle and subcutaneous adipose tissue. Diabetes 2009;58:2402-8.
20. Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zond ervan KT. Data quality control in genetic case-control association studies. Nat Protoc 2010;5:1564-73.
21. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-ba sed linkage analyses. Am J Hum Genet 2007;81:559-75.
22. Errasti-Murugarren E, Pastor-Anglada M. Drug transporter pharmacogeneti cs in nucleoside-based therapies. Pharmacogenomics 2010;11:809-41.
23. Duley JA, Somogyi AA, Martin JH. The future of thiopurine pharmacog enomics. Pharmacogenomics 2012;13:1549-52.
24. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-W heeler transform. Bioinformatics 2010;26:589-95.
25. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-ge neration DNA sequencing data. Nat Genet 2011;43:491-8.
26. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of gen etic variants from high-throughput sequencing data. Nucleic Acids Res 2 010;38:e164.
27. Gee HY, Kim YW, Jo MJ, Namkung W, Kim JY, Park HW, et al. Syna ptic scaffolding molecule binds to and regulates vasoactive intestinal poly peptide type-1 receptor in epithelial cells. Gastroenterology 2009;137:607-17, 17 e1-4.
28. Ma M, Harding HP, O'Rahilly S, Ron D, Yeo GS. Kinetic analysis of F TO (fat mass and obesity-associated) reveals that it is unlikely to functio n as a sensor for 2-oxoglutarate. Biochem J 2012;444:183-7.
29. Fischer J, Koch L, Emmerling C, Vierkotten J, Peters T, Bruning JC, et al. Inactivation of the Fto gene protects from obesity. Nature 2009;458:8 94-8.
30. Loos RJ, Yeo GS. The bigger picture of FTO: the first GWAS-identified obesity gene. Nat Rev Endocrinol 2014;10:51-61.

31. Boissel S, Reish O, Proulx K, Kawagoe-Takaki H, Sedgwick B, Yeo GS, et al. Loss-of-function mutation in the dioxygenase-encoding FTO genec auses severe growth retardation and multiple malformations. Am J Hum Genet 2009;85:106-11.
32. Rosendahl J, Witt H, Szmola R, Bhatia E, Ozsvari B, Landt O, et al. C hymotrypsin C (CTRC) variants that diminish activity or secretion are as sociated with chronic pancreatitis. Nat Genet 2008;40:78-82.
33. Fedeles BI, Singh V, Delaney JC, Li D, Essigmann JM. The AlkB Fami ly of Fe(II)/alpha-Ketoglutarate-dependent Dioxygenases: Repairing Nuclei c Acid Alkylation Damage and Beyond. J Biol Chem 2015;290:20734-42
34. Lee M, Kim B, Kim VN. Emerging roles of RNA modification: m(6)A and U-tail. Cell 2014;158:980-7.
35. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation med iated through reversible m(6)A RNA methylation. Nat Rev Genet 2014;1 5:293-306.
36. Merkestein M, McTaggart JS, Lee S, Kramer HB, McMurray F, Lafond M, et al. Changes in gene expression associated with FTO overexpressio n in mice. PLoS One 2014;9:e97162.
37. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. V ariation in FTO contributes to childhood obesity and severe adult obesity Nat Genet 2007;39:724-6.
38. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgr en CM, et al. A common variant in the FTO gene is associated with bo dy mass index and predisposes to childhood and adult obesity. Science 2 007;316:889-94.
39. Lorsbach RB, Moore J, Ang SO, Sun W, Lenny N, Downing JR. Role of RUNX1 in adult hematopoiesis: analysis of RUNX1-IRES-GFP knock-in mice reveals differential lineage expression. Blood 2004;103:2522-9.
40. Stocco G, Cheok MH, Crews KR, Dervieux T, French D, Pei D, et al. Genetic polymorphism of inosine triphosphate pyrophosphatase is a deter minant of mercaptopurine metabolism and toxicity during treatment for a cute lymphoblastic leukemia. Clin Pharmacol Ther 2009;85:164-72.
41. Krishnamurthy P, Schwab M, Takenaka K, Nachagari D, Morgan J, Lesli e M, et al. Transporter-mediated protection against thiopurine-induced he matopoietic toxicity. Cancer Res 2008;68:4983-9.

ABSTRACT (IN KOREAN)

염증성 장질환에서 Thiopurine 부작용 백혈구 감소증 원인 유전자 변이 탐색연구

< 지도교수 이 민 구>

연세대학교 대학원 의학과

김 한 상

백혈구 감소증은 면역억제제 (Thiopurine) 치료시 치명적인 부작용 중의 하나로 서양에 비해 아시아 인구에서 빈도가 흔한 것으로 알려져 있다. 본 연구는 염증성 장 질환에서 thiopurine 사용에 의한 백혈구 감소증의 원인 유전자 변이를 찾고자 하였다. 염증성 장 질환에서 thiopurine 사용 환자의 혈액 샘플을 이용하여 전장 유전체 분석을 수행하였고 후보 유전자를 대상으로 차세대 염기서열 분석을 수행하였다. 탐색된 후보 염기 변이에 대하여 2개의 다른 코호트에서 독립적으로 검증하였으며 해당 유전자에 대한 기능 연구를 수행하였다. 전장 유전체 분석을 통해 thiopurine에 의한 백혈구 감소증 관련 FTO (rs16957920) 및 RUNX1 (rs2834826) 유전자 변이를 찾을 수 있었다. 차세대 염기서열 분석을 통해 유전자 분석을 통해 FTO 단백질 변이 염기 서열변이 (rs79206939, p.A134T)가 백혈구 감소증과 관련되어 있음을 확인할 수 있었으며 추가적인 두 개의 코호트에서 이를 검증할 수 있었다. (combined $P = 1.3\times10^{-8}$, 교차비 = 4.3). FTO A134T 염기 서열 변이는 한국인에서 5.1%에 존재하며 서양에서는 0.1% 미만에서 존재하는 것으로 확인하였다. FTO의 생물학적 기능인 디메틸화 정도를 측정한

결과 FTO A134T 염기 서열 변이는 정상 단백질의 65% 정도 기능으로 감소됨을 확인할 수 있었다. FTO 유전자 제거 생쥐 모델을 이용하여 thiopurine을 투여한 결과 정상 생쥐에 비해 유전자 제거 생쥐 모델에서 thiopurine에 의한 백혈구 감소가 증가됨을 확인할 수 있었다. FTO 유전자의 A134T 염기 서열변이와 thiopurine에 의한 백혈구 감소증이 관련 있음을 확인할 수 있었다. 본 연구를 통해 밝혀진 유전자 변이를 이용하여 향후 thiopurine 치료 환자를 대상으로 백혈구 감소증 예측바이오 마커로서 활용될 수 있으리라 기대된다.

핵심되는 말: 전장 유전체 분석, 티오퓨린, 백혈구 감소증, 염증성 장 질환, FTO

PUIBLICATION LIST

1. **Kim HS***, **Cheon JH***, **Jung ES***, Park J, Aum S, Park SJ, et al. A coding variant in FTO confers susceptibility to thiopurine-induced leukopenia in East Asian patients with IBD. Gut 2017;66:1926-35.
2. **Matei I***, **Kim HS***, Lyden D. Unshielding Exosomal RNA Unleashes Tumor Growth And Metastasis. Cell 2017;170:223-5.
3. Becker A, Thakur BK, Weiss JM, **Kim HS**, Peinado H, Lyden D. Extracellular Vesicles in Cancer: Cell-to-Cell Mediators of Metastasis. Cancer Cell 2016;30:836-48.
4. **Kim HS***, **Jung M***, **Kang HN***, Kim H, Park CW, Kim SM, et al. Oncogenic BRAF fusions in mucosal melanomas activate the MAPK pathway and are sensitive to MEK/PI3K inhibition or MEK/CDK4/6 inhibition. Oncogene 2017;36:3334-45.
5. **Kim HS**, Kwon HJ, Jung I, Yun MR, Ahn MJ, Kang BW, et al. Phase II clinical and exploratory biomarker study of dacomitinib in patients with recurrent and/or metastatic squamous cell carcinoma of head and neck. Clin Cancer Res 2015;21:544-52.
6. **Yoon JK***, **Ahn J***, **Kim HS***, Han SM, Jang H, Lee MG, et al. microDuMIP: target-enrichment technique for microarray-based duplex molecular inversion probes. Nucleic Acids Res 2015;43:e28.
7. **Jung J***, **Kim HS***, Lee MG, Yang EJ, Choi JY. Novel COCH p.V123E Mutation, Causative of DFNA9 Sensorineural Hearing Loss and Vestibular Disorder, Shows Impaired Cochlin Post-Translational Cleavage and Secretion. Hum Mutat 2015;36:1168-75.
8. **Ahn JW***, **Kim HS***, **Yoon JK***, Jang H, Han SM, Eun S, et al. Identification of somatic mutations in EGFR/KRAS/ALK-negative lung adenocarcinoma in never-smokers. Genome Med 2014;6:18.