Journal of Advanced Biomedical Sciences

https://jabs.fums.ac.ir/
Online ISSN: 2783-1523

Journal of Advanced Biomedical Sciences

Fasa University of Medical Sciences

Assessing Candidate Genetic Variants Associated with Ulcerative Colitis (UC): A Case-Control Study

Mahsa Alem[1,2], Seyyed Meysam Abtahi Froushani[1], Nasim Hajighahramani[3], Saied Hosseini-Asl[2], Farhad Pourfarzi[2], Rasoul Nemati[4]

1. Department of Microbiology, Faculty of Veterinary Medicine, Urmia University, Urmia, Iran
2. Digestive Disease Research Center, Ardabil University of Medical Sciences, Ardabil, Iran
3. Department of Pharmaceutical Biotechnology, School of Pharmacy, Ardabil University of Medical Sciences, Ardabil, Iran
4. Department of Internal Medicine, Faculty of Medicine, Ardabil University of Medical Sciences, Ardabil, Iran

Article Info

Article Type:
Original Article

Article history:
Received
15 Sep 2024
Received in revised form
28 Oct 2024
Accepted
11 Jan 2025
Published online
20 Jan 2025

Publisher
Fasa University of Medical Sciences

Abstract

Background & Objectives: This study explored the immunogenetic perspective underpinning the susceptibility and progression of ulcerative colitis (UC), a form of inflammatory bowel disease (IBD) characterized by chronic auto-inflammatory processes in the gut. This multifactorial disorder occurs in genetically predisposed individuals experiencing dysbiosis, wherein genetic factors contribute to intestinal imbalance and modulate immune pathways. Recent discoveries of novel genetic loci and polymorphisms have highlighted racial diversity in IBD susceptibility. Our research addresses a gap in the literature by examining the association between specific genetic loci and UC susceptibility in a defined ethnic group, aiming to identify and evaluate candidate genetic polymorphisms as part of an immunogenetic assessment to elucidate the mechanisms underlying UC susceptibility and progression.

Materials & Methods: In this case-control study, we evaluated 794 reference sequences (rs) from global, national, and regional ethnic databases, which were utilized for genotyping 150 control subjects via blood samples. We selected and investigated five genetic variants—rs3764147 (*LACC1* or *C13orf31*), rs763780 (*IL17F*), rs3749171 (*GPR35*), rs1260326 (*GCKR*), and rs4077515 (*CARD9*)—using Amplification-refractory mutation system polymerase chain reaction (ARMs PCR) followed by Sanger sequencing for confirmation in 50 UC patients.

Results: Analysis of the candidate genetic polymorphisms revealed that the *LACC1*, *IL17F*, and *GPR35* variants were significantly associated with UC, with the primary outcome focusing on this association and the secondary outcome examining the allelic frequencies of the five variants in both groups.

Conclusion: Ultimately, our findings demonstrate that the *GPR35* (G protein-coupled receptor 35) rs3749171, *IL17F* (Interleukin 17F) rs763780, and *LACC1* (laccase domain containing 1) rs3764147 variants contribute to UC susceptibility as a polygenic trait in the studied ethnic group. This underscores the potential importance of genetic architecture across different ethnicities, offering valuable biological insights into IBD pathogenesis and guiding future research.

Keywords: Ulcerative colitis, *LACC1*, *IL17F*, *GPR35*, *CARD9*, Single Nucleotide Polymorphism (SNP)

Cite this article: Alem M, Abtahi Froushani SM, Hajighahramani N, Hosseini-Asl S, Pourfarzi F, Nemati R. Assessing Candidate Genetic Variants Associated with Ulcerative Colitis (UC): A Case-Control Study. J Adv Biomed Sci. 2025; 15(1): 77-101.
DOI: 10.18502/jabs.v15i1.17553

Corresponding Author: Seyyed Meysam Abtahi Froushani, Department of Microbiology, Faculty of Veterinary Medicine, Urmia University, Urmia, Iran.
Email: sm.abtahi@urmia.ac.ir

Introduction

Inflammatory bowel disease (IBD), which encompasses a spectrum of inflammatory conditions affecting the large and small intestines,

represents an auto-inflammatory disorder of the immune system. It typically arises in genetically predisposed individuals due to an inadequate immune response to microbial antigens derived from commensal organisms. IBD is classified as either Ulcerative Colitis (UC) or Crohn's Disease (CD), both of which primarily affect the gastrointestinal (GI) tract and are largely attributed to gut dysbiosis (1–4). The burden of IBD is considerable, with prevalence rates exceeding 0.3 percent in North America, Oceania, and most European countries. Similarly, rapidly industrializing nations in Africa, Asia, and South America are experiencing a rising incidence of IBD (5). Although IBD predominantly manifests in the GI tract, its systemic nature means it can impact virtually every organ, engaging not only gastroenterologists but also researchers from diverse fields. Moreover, IBD exemplifies the rapid evolution of biomedical knowledge and its profound impact on clinical practice and patients' quality of life (6). IBD, a chronic, relapsing, and debilitating condition, provokes inflammation in the GI tract via pro-inflammatory cytokines such as tumor necrosis factor-alpha (TNF-α), interleukin (IL)-1 beta, and IL-6 and presents with symptoms including abdominal pain, diarrhea, GI bleeding, and unintended weight loss. While UC primarily affects the colon and rectum, CD may involve any segment of the digestive system, making it a serious health concern that necessitates prompt medical intervention, continuous management, lifelong medication, and frequent hospitalizations due to the disease or its complications (6–8).

A fundamental principle in immunology is that chronic inflammation in otherwise healthy tissues can cause significant tissue damage and lead to autoimmune diseases. Therefore, it is imperative to regulate inflammatory responses to prevent long-term adverse effects (6). Although the precise immunopathogenesis of IBD remains unclear, genetic susceptibility combined with environmental risk factors plays a crucial role in triggering this multifactorial disorder. Moreover, genetic predisposition influences the immune response to infections associated with IBD (1, 9–12).

Numerous studies have investigated the impact of genetic factors on IBD development, suggesting that these factors significantly disrupt the balance of intestinal bacterial populations associated with the disease. Nevertheless, further research is needed to clarify the specific effects of genetic determinants on the principal immune pathways involved in IBD. Genetic variations may produce similar disease phenotypes, and most polymorphisms contributing to host susceptibility to IBD can be broadly categorized into those affecting immune-microbe interactions, mucosal barrier function, immunoregulation, heightened immune responses, bacterial clearance, and chronic tissue damage (9, 13, 14).

The discovery of new genetic locations (loci) related to IBD underscores the notable consideration of variations and heterogeneity between racial populations (6, 15, 16). The identification of novel genetic loci associated with IBD underscores the considerable variability and heterogeneity observed among different racial populations (6, 15, 16). Previous studies have identified the specific genetic loci associated with both CD and UC, as well as a general genetic locus for IBD. These discoveries have provided new insights into the underlying mechanisms of these diseases, including the role of autophagy.

Moreover, some of these genetic loci are also implicated in other inflammatory conditions. Extensive research has led to the identification of new genetic loci, many of which are concurrently associated with both CD and UC. Furthermore, certain genetic loci appear to have been subjected to evolutionary selective pressures, favoring either a single allele (directional selection) or the maintenance of both alleles (balancing selection) (11, 17–20). Additionally, numerous genetic loci associated with IBD are implicated in other immune-mediated disorders, such as

ankylosing spondylitis and psoriasis. There is also a significant overlap between the genetic loci predisposing individuals to IBD and those linked to mycobacterial infections.

Gene co-expression network analysis has revealed shared pathways between the host response to *Mycobacteria* and the genetic factors that predispose individuals to IBD (11, 21). In this context, the term "variant" refers to a stable alteration in the genetic sequence, distinct from a mutation. Such variants lack documented evidence linking them to clinical symptoms, and their pathogenicity has not been confirmed in reputable databases (22, 23). Redefining cytokine networks unveils the cellular heterogeneity of the immune system, emphasizing the myriad cell types that communicate via cytokines. Genes associated with IBD play a prominent role in these networks, illuminating how genetic susceptibilities impact inflammatory processes and intestinal homeostasis. Key cytokines involved include IL10RA, IL-23, IL23R, IL-22, IL-1β, IL12B, IL23A, IL-6, IL6ST, IL17, IL17F, IL-33, and IL-36 (24). Although these identified genes are intriguing, they currently lack definitive evidence to establish causality. In this study, we aim to elucidate the role of genetic polymorphisms in the UC subtype of IBD.

Specifically, we investigated the potential role of particular genetic variants in the immunopathogenesis of UC among the Azeri-Iranian population. The primary objective of this study was to identify and highlight the genes and genetic variants associated with UC in patients relative to non-UC controls. Secondary objectives included assessing the frequency and associations of these genetic variants in individuals with UC compared to controls, as well as conducting a comparative analysis of allele frequencies against existing genetic study databases. Given that polygenic diseases are not caused by a single mutation, our study involved an extensive genetic evaluation. A comprehensive assessment of gene architecture and common mutations in UC will contribute to a deeper understanding of its pathogenesis. Such insights may pave the way for future preventive strategies against this disease.

Materials and Methods

This study was conducted as a retrospective case-control investigation and is reported in accordance with the STROBE guidelines ("Strengthening the Reporting of Observational Studies in Epidemiology") (Supplementary Table 1,

Table 1. Frequency of Mutated Alleles. The frequencies of the mutated alleles for the selected gene variants were analyzed separately in the controls and cases of this study, the Azeri-Iranian population database, the Iranian population (as reported by www.iranome.com), and three established global genetic datasets (1000 Genomes, gnomAD, and ExAC). Each column represents the mutant allele frequency in each group.

Gene (rs)	Control	Case	Total	Freq Azeri	Freq Iranome	Freq World 1000 Genomes	Freq World gnomAD	Freq World ExAC
rs3764147 *LACC1*	0.256	0.44	0.302	0.320	0.296	0.306	0.267	0.270
rs763780 *IL17F*	0.036	0.12	0.057	0.055	0.056	0.093	0.065	0.067
rs3749171 *GPR35*	0.056	0.25	0.105	0.085	0.101	0.151	0.190	0.175
rs1260326 *GCKR*	0.526	0.51	0.522	0.535	0.587	0.707	0.673	0.643
rs4077515 *CARD9*	0.226	0.30	0.245	0.270	0.282	0.367	0.381	0.402

LACC1: Laccase domain containing 1, *IL-17F*: Interleukin 17F, *GPR35*: G protein-coupled receptor 35, *GCKR*: Glucokinase Regulatory Protein, *CARD9*: Caspase Recruitment Domain Family Member 9.

at the end of the article) (25). In this case-control study, after comprehensively examining the 794 reference sequences (rs) from various valid global, national, and regional ethnic databases and assessing the sequencing outcomes from genotyping control subjects, five variants were selected for laboratory investigation using the Amplification-refractory mutation system polymerase chain reaction (ARMs PCR) method with confirmation by Sanger sequencing in case individuals from the northwest region of Iran, representing the Azeri-Iranian population. Accordingly, candidate genetic variants implicated in IBD susceptibility were evaluated.

Data Collecting

A specific panel of genes was selected for examination based on an extensive literature review and established databases. Subsequently, the most influential genes were subjected to mutation analysis in the laboratory. The selection criterion for IBD-associated genes was determined by the volume of articles and documents related to each gene, as retrieved from the GAAD (Database Association Autoimmune Disease Gene) website (26) along with additional databases such as Varsome.com (27), Ensembl.org (28), Omim.org (29), SNPedia.com (30), and DisGeNET.org (31).

This study aimed to evaluate the impact of gene polymorphisms implicated in IBD, ascertain the contribution of specific variants to the disease, and examine the influence of these polymorphisms on disease susceptibility. A total of 794 rs from reputable databases were reviewed in the 150 control participants to investigate gene polymorphisms and the presence of homozygous or heterozygous variants. Among 169 missense polymorphisms, 19 rs exhibited a decreased frequency in the control group compared to global, national, and regional ethnic population studies, and these 19 rs have been reported to influence IBD. Additionally, five rs that were found to have a predisposing relationship with IBD, as reported in articles from PubMed and ScienceDirect, were selected for molecular-genomic analysis. The genes studied are involved in chronic inflammatory processes and affect the body's immune system pathways.

Sampling Method

This study utilized biosamples, specifically patient blood, collected in accordance with ethical guidelines. Blood sampling was performed only after obtaining approval from the Institutional Review Board (IRB) and the Ethical Committee of Ardabil University of Medical Sciences (IRB no.: IR.ARUMS.REC.1400.027). Furthermore, informed consent was obtained from all participants prior to blood collection (32). For genetic variant analysis, blood samples were collected from clinical participants, including 150 non-healthy, non-IBD/non-UC control patients and 50 IBD cases from the Middle East, and northwest Iran. The inclusion of non-healthy individuals as controls offers a distinct advantage, as it minimizes the potential overestimation of odds ratios—a common bias in case-control studies.

We collected samples over a two-year period, from November 2020 to October 2022. Because genetic markers remain constant over time, we employed a prevalent case selection strategy, as incident cases would not add value. A consecutive, clinic-based sampling approach was used to select 150 non-IBD and non-UC controls and 50 IBD (UC) cases, in line with the study's primary objective and empirical strategy (33–35). The non-IBD and non-UC controls were obtained from one of two provincial referral genetic laboratories. Inclusion criteria for cases required a diagnosis of IBD or UC by a gastroenterology specialist following widely accepted clinical standards. Controls were excluded if they had chronic diarrhea lasting six weeks, and both cases and controls were excluded if they had any autoinflammatory or autoimmune diseases that could confound our study. According to the gastroenterology specialist, a pathological examination was unnecessary because IBD and UC invariably present with symptoms and do not remain asymptomatic. Thus, individuals who exhibited no specific disease symptoms and had no treatment history were conclusively classified

as non-IBD. A standardized method was employed to ensure precise measurement of outcomes for both cases and controls, with all procedures conducted by a single individual in the molecular and clinical examination sections. Demographic covariates—including age and gender—were evenly distributed between the case and control groups, which consisted of white Caucasian-Persian individuals from the Ardabil region in northwest Iran. Given the influence of sample size on p-values, we refrained from testing for statistically significant differences in demographic characteristics. The average ages for the control, case, and overall groups were 21.32±8, 34.3±11, and 24.99±11 years, respectively. Regarding gender, 59.3% of controls and 62% of cases were female, with the remainder being male. Control subjects underwent genetic testing for a single nucleotide variant (SNV) associated with conditions such as deafness, which is not related to IBD, including ulcerative colitis. This testing, typically used in reproductive-age or pediatric genetic screening, was deemed appropriate for these subjects. Given that the independent variables under study are genetic and do not change over time, the lack of age matching between cases and controls is not considered a study limitation, as genetic traits are inherent across all ages.

Gene Expression Analysis

In our study, we examined 794 deleterious and associational genetic variants reported in the literature, evaluating both mutation and allele frequencies. Variants associated with IBD were identified from previous studies and curated databases. For the control group, genetic analysis was performed using whole exome sequencing (WES) (36) to detect differences in variant frequencies relative to global and local populations. Subsequently, five specific reference Single Nucleotide Polymorphisms (SNPs) (rs) were selected, and the frequencies of the mutated alleles for genetic variants in five distinct genes were compared among the control population, the Azeri population of Iran, the overall Iranian population (as provided by Iranome), and the global population using data from three studies: the 1000 Genomes Project (http://browser.1000genomes.org/index.html) (37), The Genome Aggregation Database (gnomAD) (https://registry.opendata.aws/broad-gnomad) (38), and The Exome Aggregation Consortium (ExAC) (http://exac.broadinstitute.org) (39). Details of these comparisons are summarized in Table 1.

Regarding the case subjects, following DNA extraction from whole blood samples using the sedimentation method, a one-step Master Mix was employed to perform the amplification-refractory mutation system polymerase chain reaction (ARMs PCR) assay using custom-designed primers, reagents, and precisely controlled time and temperature parameters. For each sample, the ARMs-PCR reaction was set up in a 10microliter volume for assessing *IL17F*, *GCKR*, *CARD9*, and *GPR35* gene variants—comprising 5 µl of master mix, 0.5 µl of primers, 3 µl of sterile double-distilled water, and 1.5 µl of template—and in a 15microliter volume for evaluating the *LACC1* gene variant—comprising 7.5 µl of master mix, 2.5 µl of primers, 2 µl of sterile double-distilled water, and 3 µl of template. The ARMs PCR technique targets mutations such as single nucleotide polymorphisms (SNPs) by amplifying template DNA exclusively in the presence of the target allele. Prior to processing the case samples, a gradient PCR was conducted to determine the optimal annealing temperature for the target gene variants. The thermocycler was programmed with a temperature range of 59–68°C, with 60°C in 30 cycles identified as optimal for all five variants. Subsequently, separate PCR programs were run for each gene at the specific annealing temperatures corresponding to each primer pair. The ARMs PCR protocol comprised distinct phases: primary denaturation at 95°C for 5 minutes, secondary denaturation at 94°C for 1 minute, extension at 72°C for 60 seconds, and a final extension at 72°C

Table 2. This table presents the OR results for the gene variants. ORs for *LACC1*, *IL17F*, and *GPR35* are statistically significant, as their CIs do not include 1, whereas the ORs for *GCKR* and *CARD9* are non-significant because their CIs include 1. Additionally, Pearson Chi-Square tests demonstrated significant differences between cases and controls. *Abbreviations: 95% CI (95% confidence interval); OR (Odds Ratio); UL (Upper Limit); LL (Lower Limit). Significant via Bonferroni correction.

Gene (rs)	OR	95% CI (UL to LL) (Woolf)	P-value
rs3764147 *LACC1*	2.27551	[1.419005 to 3.648999]	0.0005*
rs763780 *IL17F*	3.582645	[1.52787 to 8.400808]	0.0019*
rs3749171 *GPR35*	5.54902	[2.849014 to 10.80782]	0.0000*
rs1260326 *GCKR*	0.9354172	[0.5947428 to 1.471233]	0.7726
rs4077515 *CARD9*	1.462185	[0.8816064 to 2.425101]	0.1398

for 7 minutes. Quadruple (tetra) primers were used in the reactions, and gel electrophoresis typically revealed three bands for heterozygous samples and two bands for homozygous samples, indicating the presence of wild-type or mutant alleles based on band size. The triad (triple) primer method (Supplementary Table 2, at the end of the article) involved two separate reactions specific for wild-type and mutant allele primers, with the location of the bands indicating the corresponding alleles. Each sample underwent individual ARMs PCR reactions for the five selected gene variants, including a negative control. PCR products were loaded onto a 1.5% agarose gel to assess instrument performance, followed by electrophoresis on an acrylamide gel for polymorphism analysis (40). The ARMs PCR method was validated by Sanger sequencing (41), with 2–6% of the tetra primer strategy reactions confirming the presence of heterozygous, homozygous wild-type, and mutant samples.

Statistical Analysis

We evaluated the appropriateness of the case-control ratio, which comprised 50 case individuals and 150 control individuals (a 1:3 ratio), with an overall non-participation rate of 4–5%. Furthermore, we assessed not only the frequency of the mutated allele but also the number of individuals carrying the mutation in a heterozygous state (i.e., inheriting one wild-type and one mutated allele) versus a homozygous mutant state (i.e., inheriting two mutated alleles). The mutated allele frequency was calculated using the following formula:

Statistical analyses were performed to compare allele frequencies and genotypes between control and case subjects. Quantitative data were analyzed using the Pearson Chi-Square test, with significance set at $P \leq 0.01$ after applying the Bonferroni correction. Additionally, we employed Stata software to evaluate post-hoc power and to calculate odds ratios (OR) as measures of effect size. Post-hoc power calculations were conducted (42). Using the Woolf option in STATA 17, we computed the OR to assess the impact of genetic variants

$$\text{Allele Frequency of Case} = \frac{(\text{Homozygote Mutant Number} \times 2) + \text{Heterozygote Number}}{\textit{Case Number} \times 2}$$

$$\text{Allele Frequency of Control} = \frac{(\text{Homozygote Mutant Number} \times 2) + \text{Heterozygote Number}}{\textit{Control Number} \times 2}$$

on specific outcomes. Instead of the default method, we applied the CCI command with the Woolf option, as our sample size exceeded 30, which would otherwise have resulted in wider confidence intervals (CI). For our study—with 50 cases and 150 controls (each individual possessing two alleles per gene)—this approach was deemed more appropriate (43). Furthermore, it was unnecessary to report the frequency in each group as a CI since comparative interpretation was sufficiently provided by the OR index (44).

Results

Gene Variants Selection

We meticulously analyzed 794 reference sequences (rs) from reputable databases to identify gene variants, selecting only those with complete information. Our comprehensive analysis revealed 169 gene polymorphisms exhibiting amino acid substitutions (missense variants). Based on a literature review and the observation of lower frequencies in our 150 control participants relative to global, national, and regional populations (as shown in Table 1), we identified 19 gene polymorphisms associated with IBD (see Supplementary Table 3 at the end of the article). It is important to note that our control group comprised only healthy individuals concerning IBD, whereas the global, national, and regional datasets included both healthy and affected individuals. These 19 gene variants are believed to significantly influence IBD susceptibility. Furthermore, we selected five variants for molecular-genomic analysis in 50 case individuals, as previous studies have indicated a predisposing relationship between these variants and IBD. Notably, these genes have been reported to play key roles in IBD and to modulate immune system pathways.

Gene Variants Analysis

Our analysis focused on specific gene variants previously linked to IBD in global databases and studies. The frequencies of these variants in the study groups were determined using the ARMs PCR method and subsequently confirmed by Sanger-based sequencing. Comparing the variant frequencies in our study subjects with data from global, national, and regional databases facilitated the evaluation of their potential roles in IBD. Following DNA extraction, we examined mutations in the target genes in case participants, and ARMs PCR analysis was performed based on the band patterns observed on gel electrophoresis (Figures 1–5).

According to the statistical analysis, significant relationships with IBD were observed for all gene variants except for the last two rs. Tables 1 and 2 provide a detailed account of the statistical analysis. In this study, both the frequency of exposure (i.e., frequency of mutated alleles) and the appropriate effect size (reported as OR with its CI) were presented. Given that the prevalence of all gene variants is below 1%, we performed post hoc power calculations using STATA software rather than calculating the sample size a priori. Woolf's approximation was used to compute the CI for the OR.

Our findings revealed that the *LACC1* gene variant SNP poses a significant risk as a potential determinant factor, with a calculated power of approximately 60% (post hoc power calculation). In national, regional, and global population studies—including data from 1000 Genomes, gnomAD, and ExAC—the frequency of this gene variant was reported as 0.320, 0.296, 0.306, 0.267, and 0.270, respectively. In our study, the frequency among case and control subjects was 0.44 and 0.256, respectively, indicating a higher frequency in cases and a lower frequency in controls compared to the database populations. These results suggest that this gene variant may contribute to increased risk or a lack of protection. Additionally, we identified *IL17F*, *GPR35*, *CARD9*, and *GCKR* as secondary objectives in this study. With a study power of approximately 42%, all gene variants, except for *CARD9* and *GCKR*, exhibited significant differences between the case and control groups.

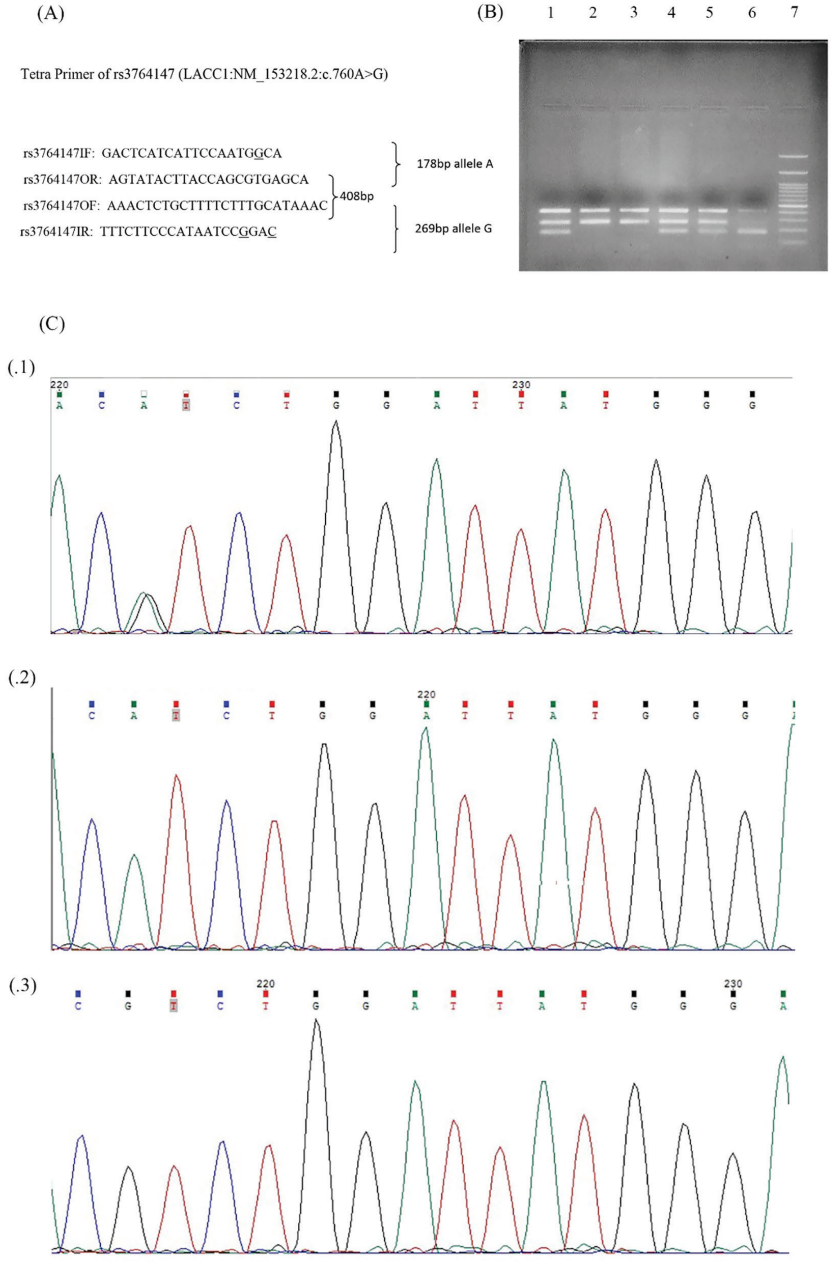


Figure 1. (A) A schematic illustrating the characteristics of the tetra primers for the *LACC1* gene variant. C (Primer = OR + OF) yields 408 bp; M (Primer = OF + IR) produces 269 bp; Wt (Primer = IF + OR) generates 178 bp. (B) An electrophoresis procedure was performed using agarose powder, 10X TAE buffer, DNA Safe Stain, and a 100 bp DNA ladder (ranging from 100 to 3000 bp). For the wells/columns, Homo M (269 bp) was observed in the 2nd and 3rd columns; Het in the 1st, 4th, and 5th columns; Homo Wt (178 bp) in the 6th column; and the ladder in the 7th (final) column. (C) Sanger sequencing was applied to confirm 6% of the tetra primer variant samples and visualization by Chromas software. (C.1) In one heterozygous sample, electrophoresis revealed three peaks, with the third peak representing both the A wild-type and mutant G alleles. (C.2) In another sample, electrophoresis of a wild-type homozygous specimen, confirmed by Sanger sequencing, displayed a second peak corresponding to the A wild-type allele. (C.3) In a homozygous mutant sample, electrophoresis and subsequent Sanger sequencing confirmed the presence of a second peak indicative of the G mutant allele. Abbreviations: Outer Reverse (OR); Outer Forward (OF); Inner Reverse (IR); Inner Forward/Common (C); Mutant (M); Wild Type (Wt); Heterozygous (Het); Tris-acetate-EDTA (TAE). Well 7: 100 Base Pair (bp) DNA marker.

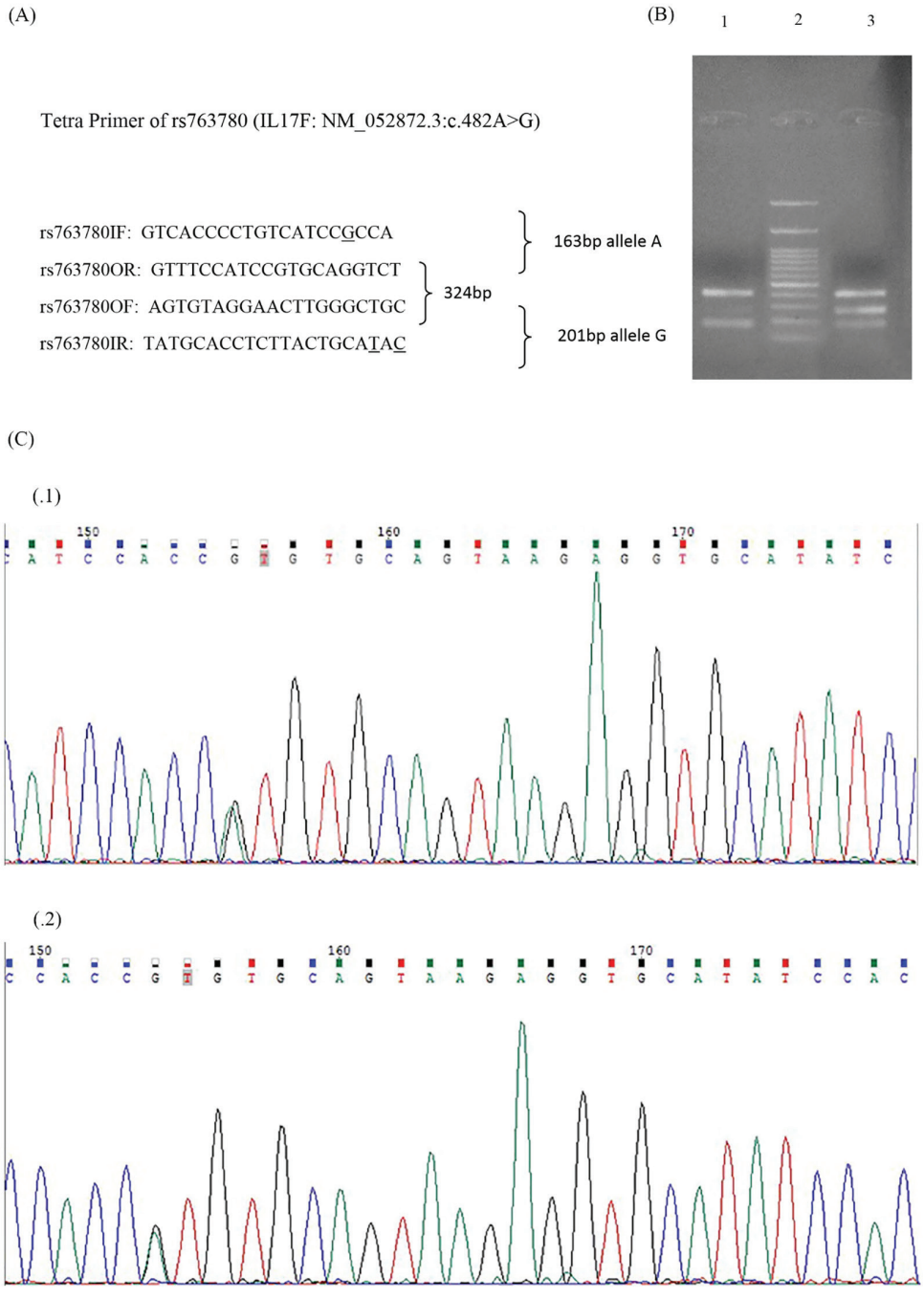


Figure 2. (A) A schematic outlining the characteristics of the tetra primers for *IL17F* gene variants. C (Primer = OR + OF) results in 324 bp; M (Primer = OF + IR) yields 201 bp; Wt (Primer = IF + OR) produces 163 bp. (B) Electrophoresis was carried out using agarose powder (Cat. No.: EP5052), 10X TAE buffer, DNA Safe Stain (Cat. No.: EP5083), and a 100 bp DNA ladder (100–3000 bp; Cat. No.: SL7041). For this gel, Homo M (201 bp) was absent in our case subjects; Het appeared in the 3rd well/column; Homo Wt (163 bp) was in the 1st well/column; and the ladder was positioned in the 2nd well/column. (C) Sanger sequencing confirmed heterozygosity in 4% of the samples, with both C.1 (8th peak) and C.2 (6th peak) aligning with the ARMs PCR results. The visualization was done by Chromas software. Abbreviations: Outer Reverse (OR), Outer Forward (OF), Inner Reverse (IR), Inner Forward/Common (C), Mutant (M), Wild Type (Wt), Heterozygous (Het), Tris-acetate-EDTA (TAE), well 2: Base Pair (bp) DNA marker.

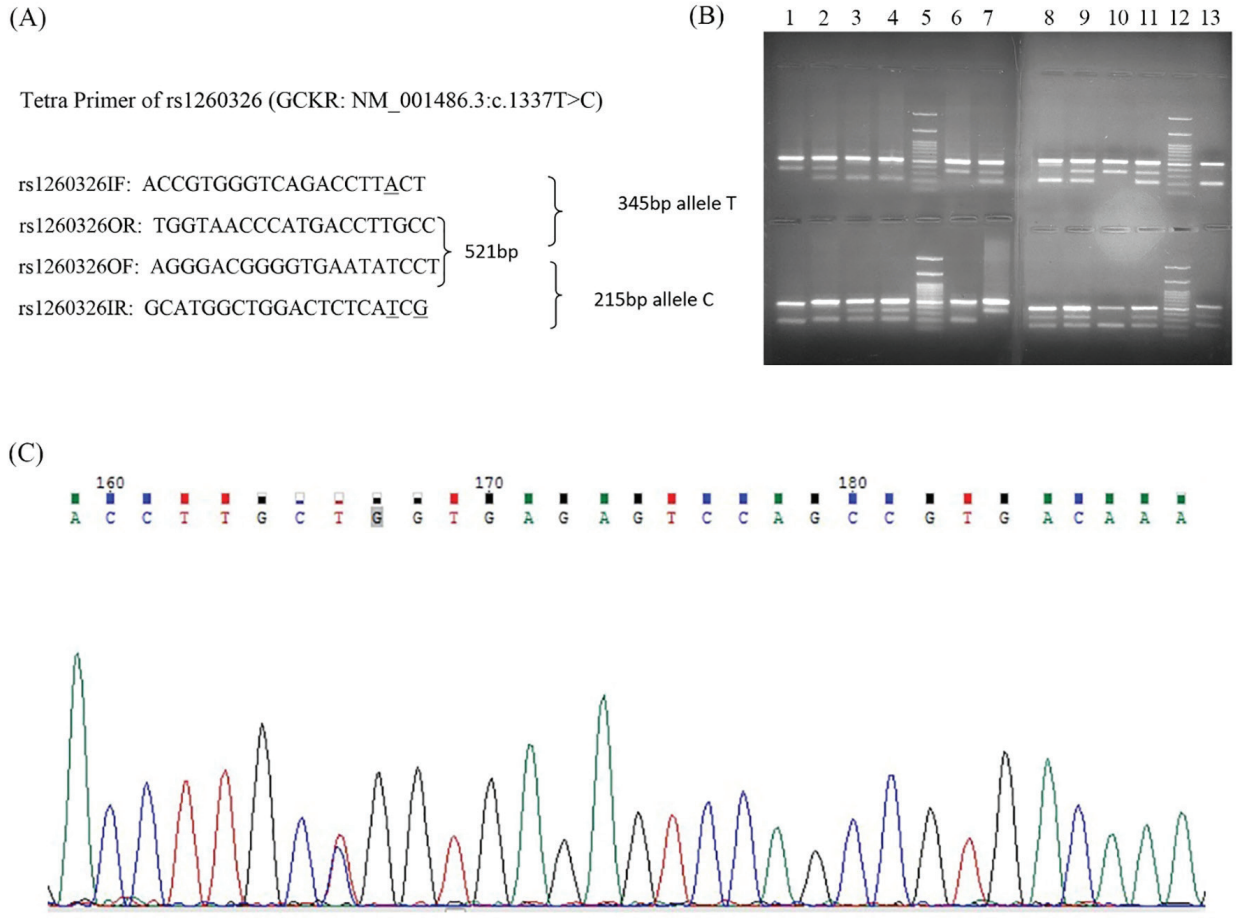


Figure 3. (A) A diagram depicting the key characteristics of the tetra primers for *GCKR* gene variants. Common (Primer = OR + OF) yields 521 bp; M (Primer = OF + IR) produces 215 bp; Wt (Primer = IF + OR) generates 345 bp. (B) Gel electrophoresis was performed using agarose powder, 10X TAE buffer, DNA Safe Stain, and a 100 bp DNA ladder (100–3000 bp). For Homo M (215 bp), the bands were observed in the 3rd, 4th, 8th, and 13th wells/columns in the top row, and the 1st, 2nd, 6th, and 10th wells/columns in the bottom row; Het was detected in the 2nd, 7th, 9th, and 11th wells/columns in the top row, and the 3rd, 4th, 8th, 9th, 11th, and 13th wells/columns in the bottom row; Homo Wt (345 bp) appeared in the 1st, 6th, and 10th wells/columns in the top row and the 7th well/column in the bottom row; the ladder was in the 5th and 12th columns. (C) In alignment with ARMs PCR, Sanger sequencing of 2% of the samples revealed heterozygosity, as evidenced by the 8th peak through Chromas software visualization. Abbreviations: Outer Reverse (OR), Outer Forward (OF), Inner Reverse (IR), Inner Forward/Common (C), Mutant (M), Wild Type (Wt), Heterozygous (Het), Tris-acetate-EDTA (TAE), wells 5 and 12: Base Pair (bp) DNA marker.

The relationship between five gene variants and IBD was analyzed using a multiple-testing approach. To ensure robust conclusions, the P-value threshold was adjusted using the Bonferroni Correction Method for both primary and secondary objectives, thereby controlling for type I error and establishing a new significance threshold. Following this correction, significant differences in gene variant frequencies between cases and controls were observed for *LACC1* (P=0.000), *IL17F* (P=0.001), and *GPR35* (P=0.000), while no significant differences were noted for *GCKR* (P=0.375) and *CARD9* (P=0.382). The primary variable—genetic variants—served as the basis for the power calculation. Among these, the primary variant of interest, *LACC1*, demonstrated a power of 60%, which, although somewhat favorable compared to the optimal power level of 80% (33–35), indicates that the study was underpowered.

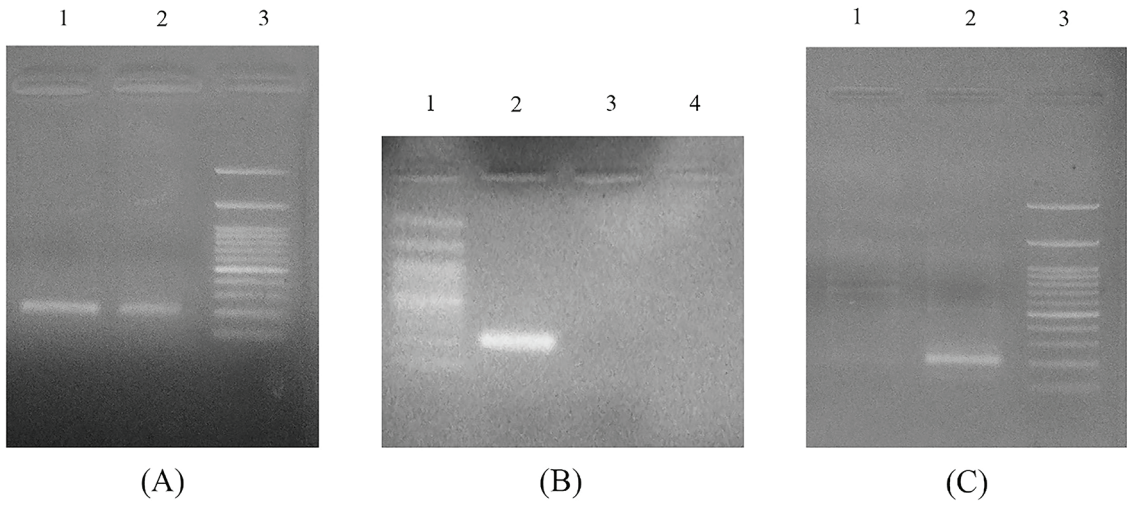
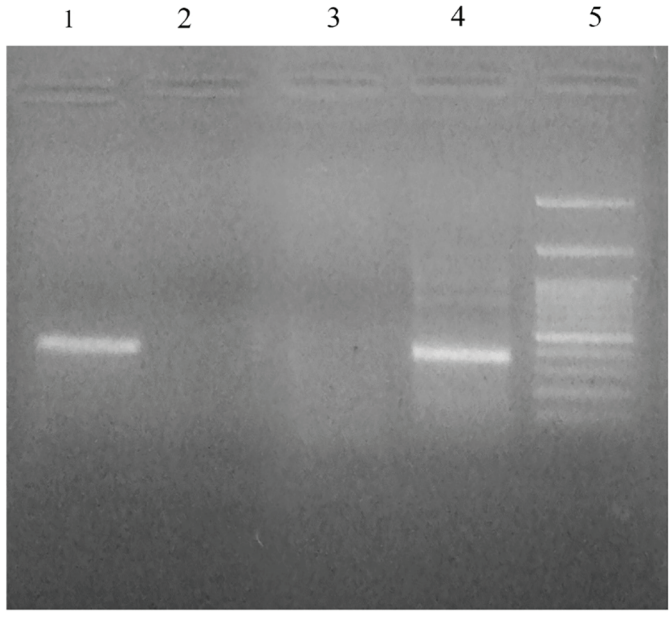


Figure 4. This figure displays agarose gel electrophoresis, conducted with 10X TAE buffer, for the *CARD9* gene variant alongside a 100 bp DNA ladder (100–3000 bp). The primers for Wt and M produced a 217 bp band. In panels (A), (B), and (C), the 3rd, 1st, and 3rd columns, respectively, represent the ladder. In summary, panels (A), (B), and (C) correspond to heterozygous, mutant, and wild-type profiles, respectively. Each image depicts different samples, with each sample loaded into two wells: the first well contained the ARMs PCR reaction with the mutant primer, and the second contained the reaction with the wild-type primer. Comparing the two wells enables the determination of the allele status for the gene variant. In panel (A), both wells display a 217 bp band, indicative of a heterozygous genotype; in panel (B), only the well with the mutant reaction shows the band; and in panel (C), only the well with the wild-type reaction shows the band. Column 4 in panel (B) represents the negative control, prepared with sterile double-distilled water. Abbreviations: Mutant (M), Wild Type (Wt), Heterozygous (Het), Tris-acetate-EDTA (TAE), wells 3A, 1B, 3C: Base Pair (bp) DNA marker.

Given that more than 95% of cases and controls participated (with only four cases and ten controls not participating), sensitivity analysis was not performed, as the low non-participation rate (less than 4–5%) was not deemed influential. Furthermore, we observed that the frequency of the TT wild-type genotype of *IL17F* (rs763780) in UC patients was lower than in the control group (76% vs. 93.33%). Analysis of *IL17F* (rs763780) polymorphisms revealed differences between the control population and IBD patients: the TT homozygous wild-type was 93.33% versus 76%, the TC heterozygous was 6% versus 24%, and the CC homozygous mutant was 0.66% versus 0%.

Discussion

In this case-control study, after a comprehensive genomic evaluation, we employed ARMs PCR—validated by Sanger sequencing for case samples and WES for controls—to genotype the selected SNPs and statistically compare allele and genotype frequencies between the groups. As the primary outcome, our research focused on five genetic variants—namely, *LACC1* (rs3764147), *IL17F* (rs763780), *GPR35* (rs3749171), *GCKR* (rs1260326), and *CARD9* (rs4077515)—in IBD patients compared to non-healthy, non-IBD controls, thereby assessing susceptibility to IBD. The secondary objectives involved comparing allele frequencies between the two groups, as well as against data from global, national, and regional ethnic studies obtained from genetic databases. A comprehensive genetic evaluation was conducted, acknowledging that a single genetic mutation does not solely cause IBD. This broader molecular understanding of the disease holds promise for developing more targeted therapeutic strategies.



Triad Primer of rs3749171 (GPR35: NM_005301.3:c.323C>T)

Figure 5. This figure presents agarose gel electrophoresis results for the *GPR35* gene variant, using 10X TAE buffer and a 100 bp DNA ladder (100–3000 bp). The primers for Wt and M yielded a 390 bp band, with the 5th column representing the ladder. In columns 1 and 2, a single sample's reaction products were loaded: the first column (mutant primer) showed a 390 bp band, while the second column (wild primer) did not, indicating the presence of the mutant allele. In columns 3 and 4, another sample exhibited a 390 bp band only in the second column (wild primer), with the first column (mutant primer) lacking a band, thereby indicating the presence of the wild-type allele. Abbreviations: Mutant (M), Wild Type (Wt), Heterozygous (Het), Tris-acetate-EDTA (TAE), well 5: Base Pair (bp) DNA marker.

Our findings revealed that individuals carrying the mutated *LACC1* allele had 2.3 times higher odds of developing IBD compared to those with the wild-type allele. Similarly, mutated *IL17F* alleles were associated with 3.6 times higher odds, and mutations in *GPR35* conferred 5.5 times higher odds. For *CARD9*, the odds were 50% higher in mutated individuals, whereas *GCKR* mutations correlated with a 7% lower risk. According to Olivier et al.'s categorization, with a 25% prevalence of IBD in our study, LACC1 (OR=2.27) and IL17F (OR=3.6) exhibited medium-strength associations, *CARD9* (OR=1.5) demonstrated a small association, and *GPR35* (OR=5.5) showed a substantial association; in contrast, *GCKR*'s OR of 0.9 indicated a trivial inverse relationship. Given that 1/1.31 approximates 0.76, the range from 0.76 to 1.30 is considered trivial (45). The association between *LACC1* (rs3764147) on chromosome 13 and Crohn's Disease aligns with previous research, while associations between UC and both *GPR35* (rs3749171) and *CARD9* (rs10781499) on chromosomes 2 and 9, respectively, were also noted. Although our results confirmed significant associations with *LACC1* and *GPR35*, *CARD9* exhibited only a negligible effect (11).

Many loci implicated in IBD remain poorly understood, with many of their functions still unclear; for instance, *C13orf31* on chromosome 13q14—which harbors the *LACC1* gene—plays a role in modulating immune responses (46). LACC1 influences innate immunity by modulating NOD2 signaling and enhancing pattern recognition receptor (PRR) activity,

although the risk variant appears to diminish PRR function. PRR functions include generating mitochondrial reactive oxygen species (ROS), bacterial clearance, signal transduction, and cytokine release. *LACC1* genetic variants have been linked to reduced PRR-induced responses and are associated with several immune-mediated diseases, such as an increased risk of Crohn's Disease (19, 47). The *LACC1* Ile254Val polymorphism is associated with Crohn's Disease and leprosy (19, 48), as well as with both Crohn's Disease and ankylosing spondylitis (AS) (19, 20). Additionally, a rare loss-of-function variant in *LACC1* (Cys284Arg) has been associated with early-onset Crohn's Disease and juvenile arthritis, exhibiting reduced ROS production (48). Furthermore, studies have demonstrated a link between this genetic variant, Crohn's Disease development, and leprosy caused by *Mycobacterium leprae* infection (48, 49). These findings suggest a shared immunopathogenesis for IBD and other inflammatory-immunological diseases, implying that common deleterious genetic polymorphisms may underlie these conditions (20). Research indicates that different populations may exhibit distinct genetic risk profiles for complex diseases like IBD (49). In our study, the missense variant rs3764147 of *LACC1* was significantly associated with IBD (OR=2.27, 95% CI: 1.42–3.65), and a positive incremental effect was observed, particularly in UC. This finding is consistent with previous studies reporting similar associations in various populations, including Europeans with Crohn's Disease (OR=1.19, 95% CI: 1.08–1.31) (20,48), Northeast Europeans with UC (OR=2.44) (49), and Caucasians from Canada and the USA (OR=1.23 for Crohn's Disease) (46). *LACC1* encodes a protein involved in fatty acid metabolism and inflammation, with altered expression observed in IBD patients (48). A specific SNP in *LACC1* is linked to a key regulator of immunometabolic function, although the functional impact of the rs3764147 variant and its interaction with environmental factors remains to be elucidated.

Conversely, our findings revealed a significant association between the *IL17F* 7488T/C variant and the odds of UC (OR=3.6, 95% CI: 1.53–8.40). This result is consistent with previous studies reporting similar associations in Japanese (UC, OR=1.85, 95% CI: 1.06–3.22) (50) and Chinese populations (CD, OR=1.18, 95% CI: 1.41–3.04), as well as with clinical parameters of both UC and CD (51), although it contrasts with findings in Caucasian populations—such as those from Germany (52) and Portugal (CD, 53)—and in Chinese Asians, where this polymorphism appears to confer only weak protection against UC (54). Moreover, a meta-analysis has demonstrated strong links between rs763780 and other inflammatory illnesses, but not with IBD (55). These discrepancies may be attributed to differences in race, geography, and sample size (51). Our study also uncovered significant differences in genotypic frequencies between UC patients and controls, a finding that aligns with some reports (51) yet diverges from others (54), likely reflecting the diverse ethnic composition of Iran. Specifically, we observed a decrease in the wild-type allele and a corresponding increase in the heterozygous allele among UC patients. Given that IL17F, secreted by T helper 17 (Th17) cells, plays a critical role in mucosal immunity, granulopoiesis, neutrophil mobilization, and host defense against pathogens (51, 56–58), the rs763780 variant—which results in a non-synonymous His161Arg substitution—may alter its function, receptor binding, secretion, and overall susceptibility to inflammatory disease (51, 59, 60), although its effects appear to vary across populations. For instance, the frequency of rs763780 is notably lower in European and American populations (59).

The *GPR35* rs3749171 variant was significantly associated with UC, exhibiting an OR of 5.55 (95% CI: 2.85–10.81), a finding that corroborates previous research (17,18). *GPR35*

encodes a G protein-coupled receptor expressed in intestinal epithelial and immune cells, where it participates in inflammation, immune regulation, and barrier function (17, 61). Acting as a receptor for kynurenic acid within the tryptophan metabolic pathway, GPR35 may prolong inflammatory conditions, as elevated kynurenic acid levels have been observed in IBD patients (17, 61–63). The rs3749171 variant induces a missense mutation that replaces threonine with methionine, potentially impacting GPR35 signaling. Genome-wide association studies (GWAS) have identified this variant as a risk locus, particularly in immune-mediated diseases such as primary sclerosing cholangitis (PSC) and UC, thereby suggesting shared pathogenic mechanisms (17). Furthermore, research on chronic inflammatory diseases—including AS, CD, psoriasis, PSC, and UC—has highlighted rs3749171 as a significant susceptibility locus in European populations, indicating a genetic distinction between PSC and classical IBD phenotypes (18).

In contrast, we were unable to replicate the previously reported positive association of the *GCKR* rs1260326 missense variant with IBD risk; both the chi-square test and the OR of 0.9 were non-significant. This finding contrasts with a recent cohort study that reported an association between the altered allele C of rs1260326 and reduced intestinal inflammation in allogeneic hematopoietic cell transplantation recipients with gut graft-versus-host disease (GVHD), a condition sharing certain features with IBD. Given that *GCKR* encodes a glucokinase regulator involved in glucose metabolism and inflammation—and that its expression is altered in IBD—the precise role of this gene in IBD pathogenesis warrants further investigation (64).

Previous research has indicated that *CARD9* rs4077515 is associated with CD and exhibits a positive additive effect (+/↑CD) (21, 65, 66); however, our study found no significant association between the *CARD9* rs4077515 variant and UC (OR=1.5, 95% CI: 0.88–2.42), possibly due to low statistical power and differences in patient populations. This result diverges from earlier studies that suggest a protective effect of the *CARD9* rs4077515 allele C against inflammatory autoimmune disorders (67), which may be explained by our exclusive focus on UC patients and the variant's potential to reduce *CARD9* expression. Since *CARD9* encodes a proinflammatory protein involved in the immune response to infections via cytoplasmic pattern recognition receptors such as NOD2—and mediates antigen presentation through dendritic cell maturation (66, 67)—alterations in *CARD9* expression have been documented in IBD patients (65–67). The rs4077515 variant (Ser12Asn) may influence *CARD9* expression, its aberrant activation, degradation, and subsequent immune interactions (67). Furthermore, previous studies have reported associations between CARD9 and NOD2 with CD, underscoring the importance of the bacterial cell wall peptidoglycan response pathway (66). GWAS have also highlighted *CARD9* as a significant susceptibility locus for CD via IL18RAP (21, 65) and for UC in European populations (68), particularly among the Dutch (17, 65), with familial interactions influencing immune-related diseases (17). Additional research is necessary to elucidate the functional impact of the *CARD9* variant on protein activity and its interactions with the microbiota.

The absence of an association between *GCKR* and *CARD9* variants and IBD in the Azeri-Iranian population suggests that these specific genes may not contribute significantly to pathogenesis in this demographic, or that their effects are obscured by other genetic or environmental factors. It is conceivable that these variants play distinct roles in different populations or disease phenotypes, thereby warranting further investigation in larger, more diverse cohorts.

In examining the links between genetic variants

and the risk of developing IBD, it is essential to consider potential sources of heterogeneity—including participants' ethnicity, lifestyle, family history, demographic characteristics, environmental exposures, and the intrinsic variability of IBD. Addressing these factors requires the use of prospective cohort studies or modern case-control methodologies, such as case-cohort and nested case-control designs, with sufficient statistical power to elucidate the true relationship. Thus, the interpretation of our findings should be approached with caution, and comprehensive prospective multicenter studies are imperative to validate the predictive effects of genetic polymorphisms on IBD risk.

Research has demonstrated that employing diverse ancestral populations in GWAS effectively identifies novel risk polymorphisms for IBD (70). Although genetic variants can significantly impact gene function, they may be too rare to detect using standard GWAS methods in IBD patients; advanced sequencing techniques may enhance their detection (71). Studies of IBD across varied ancestries—including European, East Asian, Indian, and Iranian populations—reveal shared genetic risks and highlight differences in allele frequencies and effect sizes that contribute to genetic heterogeneity. Such ancestry linkage studies are fundamental for mapping loci associated with complex diseases like IBD and for understanding their genetic architecture (16, 46, 67, 69). Moreover, integrating SNP analysis with pathway-based approaches is crucial for evaluating genetic risk (66). It is well recognized that immune-related diseases share many genetic susceptibility loci—a phenomenon known as pleiotropy (18, 20, 72)—which indicates potential therapeutic overlaps among these conditions (17).

To advance future research endeavors, it is imperative to replicate and conduct comparative analyses of findings across diverse ethnic groups and regions. This strategy will facilitate the identification of genetic diversity and population-specific factors that contribute to IBD susceptibility and heterogeneity, bolster the generalizability and validity of genetic associations, and uncover potential gene-environment interactions that modulate IBD risk (1). Furthermore, exploring these findings within various IBD subtypes and phenotypes—such as Crohn's Disease, Ulcerative Colitis, ileocolonic involvement, structuring behavior, and extraintestinal manifestations—holds promise for elucidating the complexity of the disease and for identifying gene-gene interactions that influence clinical outcomes (37, 73). Such research may also illuminate the evolutionary and historical origins of IBD and its genetic variants. Understanding the functional mechanisms and environmental interactions of these genetic variants in cell culture, animal models, and human tissues will provide critical insights into the molecular pathways and immunological processes underlying IBD development and progression. Ultimately, this line of inquiry may yield novel therapeutic targets and strategies for preventing and treating the disease. Moreover, integrating these findings with data from transcriptomics, proteomics, metabolomics, epigenomics, and microbiomics may reveal the multi-level and multi-scale interactions that shape the pathophysiology and phenotype of IBD. Finally, translating these discoveries into clinical practice and public health policy holds the potential to alleviate the burden and cost of IBD, thereby enhancing the quality of life for patients and their families (1, 11, 18, 50, 53, 55, 71, 72, 74–77).

We acknowledge several limitations in this study. First, the use of a consecutive clinic-based catchment area as a non-random sample introduces inherent biases, as clinic-based secondary sources are methodologically less robust and offer poorer generalizability than random sampling. Additionally, consecutive case and control sampling may induce selection bias. Second, a post hoc power calculation

was employed in lieu of an a priori sample size estimation due to the fixed sample size from the Unicenter study and the resulting low-power status. Despite nearly two years of patient recruitment, the low frequency of mutated alleles in each gene variant contributed to the study's low power. While many genetic and immunogenetic studies are characterized by low power, this issue is not unique to our research, underscoring the need for larger sample sizes in future investigations to enable a more comprehensive analysis; secondary or tertiary systematic reviews are also recommended to enhance certainty. Third, the absence of socioeconomic status as a demographic variable represents another limitation. Fourth, although environmental factors also contribute to disease pathogenesis, our study's exclusive focus on genetics may be viewed as a further constraint. Fifth, individual matching was not performed; instead, group matching was applied to control for confounders (35, 78, 79), with only age and gender reported for both groups. Sixth, despite employing consistent measurement methods for both cases and controls, the molecular assays differed—WES was utilized for controls, whereas cases were analyzed using ARMs PCR, with only 10% of case samples confirmed via Sanger sequencing. Although WES is considered the gold standard, its high cost necessitated a more economical approach for cases. Seventh, ARMs PCR is limited by its non-automated nature, lower precision and sensitivity, reliance on size-based discrimination prone to human error, and the absence of standardized protocols for research. Eighth, the overestimation bias associated with the OR was apparent, as our data indicated overestimation exceeding 10%. Given that the outcome prevalence was approximately 25% (50/200), the OR does not approximate the relative risk (RR); in cohort studies, the point estimate for *LACC1* might be around 1.9 rather than 2.3, suggesting that such designs would yield more conclusive causal evidence. Ninth, further subgroup analyses—stratified by age, gender, and variant allelic groups—are warranted for future studies.

Nonetheless, we believe that this study has provided valuable insights into genetic variation in IBD within specific ethnic groups, thereby paving the way for a deeper understanding of the complex genetic relationships underlying IBD pathogenesis. It is important to note that the retrospective case-control design was chosen for its exploratory rather than confirmatory nature. Considering the inherent stability of genetic variables over time, the use of a prevalent case selection pattern was appropriate, as incident cases would not have added further value, and thus this design should not be viewed as a weakness. Moreover, sampling was conducted at a specialized medical clinic, and all measurements—including genetic and clinical assessments—were performed by a single individual using specialist medical records to minimize information bias. The tools employed, including Unicenter, were highly reliable, calibrated, and administered by experts, thereby ensuring the validity of the study.

Moreover, our results must be interpreted with caution and validated in larger, more diverse cohorts and applied studies. Because genetic studies conducted in cohort settings do not inherently establish causality, future research should examine the relationship between genetic variants and diseases such as IBD within the framework of cohort and prospective studies. The value of cohort and prospective studies in this context can be summarized in two key points: First, their larger sample sizes reduce random error and yield narrower confidence intervals for effect size estimates, thereby enhancing statistical strength and evidentiary value. Second, they provide more valid and accurate measurements of not only genetic exposures, but also outcome variables and confounders, due to higher validity and reliability in measurement. For example,

cohort outcomes are based on incident (new) cases, whereas cross-sectional and case-control studies combine both old and new cases, which can diminish their value. Furthermore, cohorts typically yield more accurate measurements of confounding variables (79–81). Although the identified genes are of considerable interest, conclusive evidence establishing causality remains lacking. Therefore, it is advisable to conduct further research—particularly well-defined cohort studies, nested case-control designs, or case-cohort studies with large sample sizes using reliable and accurate genotyping methods—while adjusting for potentially confounding demographic factors, epigenetics, and lifestyle factors such as smoking (1, 49, 66, 69, 73, 74, 82). Additionally, systematic reviews, meta-analyses, and umbrella reviews are necessary to enhance statistical power for detecting relatively small associations between genetic factors and complex diseases when pooling data from individual studies, thereby clarifying the precise contribution of gene network relationships in IBD and achieving conclusive evidence of causality. Secondary and tertiary review studies also represent the strongest evidence for the direct association of genetic variants with overall IBD susceptibility across different ethnicities. Future investigations should focus on the interaction between genetic variants and environmental factors, which will further elucidate the functional mechanisms underlying IBD pathogenesis.

Conclusion

This study presents an experimental evaluation of genetic variants in relation to overall susceptibility and risk of UC as a form of IBD. In summary, our investigation assessed the effects of the *LACC1* rs3764147, *IL17F* rs763780, and *GPR35* rs3749171 variants on UC susceptibility in a specific Iranian ethnic group, while suggesting a lack of association for the *GCKR* rs1260326 and *CARD9* rs4077515 variants. Our findings provide new insights into the underlying genetic factors contributing to UC susceptibility. A comprehensive evaluation of the genetic basis of IBD will enhance our understanding of its pathogenesis as a polygenic trait, potentially paving the way for advancements in the prevention, prognosis, and treatment of this increasingly prevalent condition.

Acknowledgment

Urmia University and Ardabil University of Medical Sciences supported this work.

Conflict of Interest

The authors declare no commercial or financial conflicts of interest in the research.

Funding

The authors declare they did not obtain any dedicated funding for this project.

Ethical Considerations

This study utilized patient blood samples, and we ensured all procedures were ethically sound by securing approval from the IRB and Ethical Committee, as indicated by the specified code.

Code of Ethics

The approval code (IRB no.: IR.ARUMS.REC.1400.027) was obtained from the "Institutional Review Board" (IRB) and the Ethical and Committee approval of Ardabil University of Medical Sciences.

Authors' Contribution

Conceptualization: MA, SMAF, SH-A, FP; Methodology: MA, SMAF, SH-A. FP, NH, RN; Formal analysis and investigation: MA, SH-A, FP; Writing - original draft preparation: MA, SMAF, SH-A; Writing - review and editing: All of the authors; Funding acquisition: SMAF; Resources: SMAF, SH-AFP, RN; Supervision: SMAF

References

1 Ananthakrishnan AN. Epidemiology and risk factors for IBD. Nat Rev Gastroenterol Hepatol. 2015; 12(4):205–17.
2 Verma A, Sharda S, Rathi B, Somvanshi P, Pandey BD. Elucidating potential molecular signatures through host-microbe interactions for reactive arthritis and inflammatory bowel disease using combinatorial approach. Sci Rep. 2020; 10(1):15131.
3 Cohen-Mekelburg S, Johnson J, Paine E, Prasad MA, Dominitz JA, Hou J. Assessment of Physician Needs and Access to Inflammatory Bowel Disease Specialty Care Resources in a National Integrated Health System. Dig Dis Sci. 2024; 69(9):3180–7.
4 Gray C, Shakir R, Tumin D, Mandelia C. Predictors of Unplanned Health Care Utilization Among Children with Inflammatory Bowel Disease in a Rural Region of the Southeastern US. Dig Dis Sci. 2024; 69: 4347–54.
5 Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet. 2017; 390(10114):2769–78.
6 Baumgart DC, editor. Crohn's Disease and Ulcerative Colitis: From Epidemiology and Immunobiology to a Rational Diagnostic and Therapeutic Approach. Cham: Springer International Publishing; 2017. http://link.springer.com/10.1007/978-3-319-33703-6
7 Yeom JE, Kim SK, Park SY. Regulation of the Gut Microbiota and Inflammation by β-Caryophyllene Extracted from Cloves in a Dextran Sulfate Sodium-Induced Colitis Mouse Model. Molecules. 2022; 27(22):7782.
8 Pithadia AB, Jain S. Treatment of inflammatory bowel disease (IBD). Pharmacol Rep. 2011; 63(3):629–42.
9 Sartor RB. Microbial influences in inflammatory bowel diseases. Gastroenterology. 2008; 134(2):577–94.
10 Mann EA, Saeed SA. Gastrointestinal infection as a trigger for inflammatory bowel disease. Curr Opin Gastroenterol. 2012; 28(1):24–9.
11 Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, et al. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature. 2012; 491(7422):119–24.
12 Cao SS. Cellular Stress Responses and Gut Microbiota in Inflammatory Bowel Disease. Gastroenterol Res Pract. 2018; 2018:7192646.
13 Sartor RB. Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. Nat Clin Pract Gastroenterol Hepatol. 2006; 3(7):390–407.
14 Hu X, Yang F. Analysis of the Therapeutic Effect of Changyanning on Intestinal Flora in Inflammatory Bowel Disease. Contrast Media & Molecular Imaging. 2022 ;2022(1):3757763.
15 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(5):596–604.
16 Hong M, Ye BD, Yang SK, Jung S, Lee HS, Kim BM, et al. Immunochip Meta-Analysis of Inflammatory Bowel Disease Identifies Three Novel Loci and Four Novel Associations in Previously Reported Loci. J Crohns Colitis. 2018; 12(6):730–41.
17 Ellinghaus D, Folseraas T, Holm K, Ellinghaus E, Melum E, Balschun T, et al. Genome-wide association analysis in primary sclerosing cholangitis and ulcerative colitis identifies risk loci at GPR35 and TCF4. Hepatology. 2013; 58(3):1074–83.
18 Ellinghaus D, Jostins L, Spain SL, Cortes A, Bethune J, Han B, et al. Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nat Genet. 2016; 48(5):510–8.
19 Lahiri A, Hedl M, Yan J, Abraham C. Human LACC1 increases innate receptor-induced responses and a LACC1 disease-risk variant modulates these outcomes. Nat Commun. 2017; 8:15614.
20 Danoy P, Pryce K, Hadler J, Bradbury LA, Farrar C, Pointon J, et al. Association of Variants at 1q32 and STAT3 with Ankylosing Spondylitis Suggests Genetic Overlap with Crohn's Disease. PLoS Genet. 2010; 6(12):e1001195.
21 Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, Ahmad T, et al. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet. 2010; 42(12):1118–25.
22 Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine. 2015; 17(5):405–24.
23 Ipe J, Swart M, Burgess K, Skaar T. High-Throughput Assays to Assess the Functional Impact of Genetic Variants: A Road Towards Genomic-Driven Medicine. Clinical and Translational Science. 2017; 10(2):67–77.

24 Graham DB, Xavier RJ. Pathway paradigms revealed from the genetics of inflammatory bowel disease. Nature. 2020; 578(7796):527–39.
25 von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007; 4(10):e296.
26 Lu G, Hao X, Chen WH, Mu S. GAAD: A Gene and Autoimmiune Disease Association Database. Genomics Proteomics Bioinformatics. 2018; 16(4):252–61.
27 Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Albarca Aguilera M, Meyer R, et al. VarSome: the human genomic variant search engine. Bioinformatics. 2019; 35(11):1978–80.
28 Martin FJ, Amode MR, Aneja A, Austine-Orimoloye O, Azov AG, Barnes I, et al. Ensembl 2023. Nucleic Acids Research. 2023; 51(D1):D933–41.
29 Amberger JS, Bocchini CA, Scott AF, Hamosh A. OMIM.org: leveraging knowledge across phenotype-gene relationships. Nucleic Acids Res. 2019; 47(D1):D1038–43.
30 Cariaso M, Lennon G. SNPedia: a wiki supporting personal genome annotation, interpretation and analysis. Nucleic Acids Res. 2012; 40(Database issue):D1308-1312.
31 Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res. 2020; 48(D1):D845–55.
32 Mehta P, Zimba O, Gasparyan AY, Seiil B, Yessirkepov M. Ethics Committees: Structure, Roles, and Issues. J Korean Med Sci. 2023; 38(25):e198.
33 Van Meter E, Charnigo R. Strengthening Interactions between Statisticians and Collaborators: Objectives and Sample Sizes. J Biom Biostat. 2014; 5(1):e127.
34 Das S, Mitra K, Mandal M. Sample size calculation: Basic principles. Indian J Anaesth. 2016; 60(9):652–6.
35 Kline, R. B. Chapter 6. Statistics Reform. In: Becoming a Behavioral Science Researcher: Second Edition: A Guide to Producing Research That Matters. 2nd ed. Guilford Press; 2020. p. 118–14. https://www.guilford.com/books/Becoming-a-Behavioral-Science-Researcher/Rex-Kline/9781462538799
36 Eghbali M, Fatemi KS, Salehpour S, Abiri M, Saei H, Talebi S, et al. Whole-Exome Sequencing Uncovers Novel Causative Variants and Additional Findings in Three Patients Affected by Glycogen Storage Disease Type VI and Fanconi–Bickel Syndrome. Front Genet. 2021; 11:601566.
37 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015; 526(7571):68–74.
38 Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, et al. A genomic mutational constraint map using variation in 76,156 human genomes. Nature. 2024; 625(7993):92–100.
39 Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res. 2017; 45(D1):D840–5.
40 Little S. Amplification-refractory mutation system (ARMS) analysis of point mutations. Curr Protoc Hum Genet. 2001; 9.8:1–12.
41 Crossley BM, Bai J, Glaser A, Maes R, Porter E, Killian ML, et al. Guidelines for Sanger sequencing and molecular assay monitoring. J Vet Diagn Invest. 2020; 32(6):767–75.
42 Hickey GL, Grant SW, Dunning J, Siepe M. Statistical primer: sample size and power calculations-why, when and how? Eur J Cardiothorac Surg. 2018; 54(1):4–9.
43 Woolf B. On estimating the relation between blood group and disease. Ann Hum Genet. 1955; 19(4):251–3.
44 Dziak JJ, Dierker LC, Abar B. The Interpretation of Statistical Power after the Data have been Gathered. Curr Psychol. 2020; 39(3):870–7.
45 Olivier J, May WL, Bell ML. Relative effect sizes for measures of risk. Communications in Statistics - Theory and Methods. 2017; 46(14):6774–81.
46 Kevans D, Silverberg MS, Borowski K, Griffiths A, Xu W, Onay V, et al. IBD Genetic Risk Profile in Healthy First-Degree Relatives of Crohn's Disease Patients. J Crohns Colitis. 2016; 10(2):209–15.
47 Huang C, Hedl M, Ranjan K, Abraham C. LACC1 Required for NOD2-Induced, ER Stress-Mediated Innate Immune Outcomes in Human Macrophages and LACC1 Risk Variants Modulate These Outcomes. Cell Rep. 2019; 29(13):4525-4539.e4.
48 Cader MZ, Boroviak K, Zhang Q, Assadi G, Kempster SL, Sewell GW, et al. C13orf31 (FAMIN) is a central regulator of immunometabolic function. Nat Immunol. 2016; 17(9):1046–56.
49 Skieceviciene J, Kiudelis G, Ellinghaus E, Balschun T, Jonaitis LV, Zvirbliene A, et al. Replication study of ulcerative colitis risk loci in a Lithuanian-Latvian

case-control sample. Inflamm Bowel Dis. 2013; 19(11):2349–55.
50 Arisawa T, Tahara T, Shibata T, Nagasaka M, Nakamura M, Kamiya Y, et al. The influence of polymorphisms of interleukin-17A and interleukin-17F genes on the susceptibility to ulcerative colitis. J Clin Immunol. 2008; 28(1):44–9.
51 Zhang X, Yu P, Wang Y, Jiang W, Shen F, Wang Y, et al. Genetic polymorphisms of interleukin 17A and interleukin 17F and their association with inflammatory bowel disease in a Chinese Han population. Inflamm Res. 2013; 62(8):743–50.
52 Seiderer J, Elben I, Diegelmann J, Glas J, Stallhofer J, Tillack C, et al. Role of the novel Th17 cytokine IL-17F in inflammatory bowel disease (IBD): upregulated colonic IL-17F expression in active Crohn's disease and analysis of the IL17F p.His161Arg polymorphism in IBD. Inflamm Bowel Dis. 2008; 14(4):437–45.
53 Durães C, Machado JC, Portela F, Rodrigues S, Lago P, Cravo M, et al. Phenotype-genotype profiles in Crohn's disease predicted by genetic markers in autophagy-related genes (GOIA study II). Inflamm Bowel Dis. 2013; 19(2):230–9.
54 Chen B, Zeng Z, Hou J, Chen M, Gao X, Hu P. Association of interleukin-17F 7488 single nucleotide polymorphism and inflammatory bowel disease in the Chinese population. Scand J Gastroenterol. 2009; 44(6):720–6.
55 Eskandari-Nasab E, Moghadampour M, Tahmasebi A. Meta-Analysis of Risk Association Between Interleukin-17A and F Gene Polymorphisms and Inflammatory Diseases. J Interferon Cytokine Res. 2017; 37(4):165–74.
56 Fujino S, Andoh A, Bamba S, Ogawa A, Hata K, Araki Y, et al. Increased expression of interleukin 17 in inflammatory bowel disease. Gut. 2003; 52(1):65–70.
57 Wright JF, Guo Y, Quazi A, Luxenberg DP, Bennett F, Ross JF, et al. Identification of an interleukin 17F/17A heterodimer in activated human CD4+ T cells. J Biol Chem. 2007; 282(18):13447–55.
58 Kobayashi T, Okamoto S, Hisamatsu T, Kamada N, Chinen H, Saito R, et al. IL23 differentially regulates the Th1/Th17 balance in ulcerative colitis and Crohn's disease. Gut. 2008; 57(12):1682–9.
59 Kawaguchi M, Takahashi D, Hizawa N, Suzuki S, Matsukura S, Kokubu F, et al. IL-17F sequence variant (His161Arg) is associated with protection against asthma and antagonizes wild-type IL-17F activity. J Allergy Clin Immunol. 2006; 117(4):795–801.
60 Zacarias JMV, Sippert EÂ, Tsuneto PY, Visentainer JEL, de Oliveira e Silva C, Sell AM. The Influence of Interleukin 17A and IL17F Polymorphisms on Chronic Periodontitis Disease in Brazilian Patients. Mediators Inflamm. 2015; 2015:147056.
61 Wang J, Simonavicius N, Wu X, Swaminath G, Reagan J, Tian H, et al. Kynurenic acid as a ligand for orphan G protein-coupled receptor GPR35. J Biol Chem. 2006; 281(31):22021–8.
62 Forrest CM, Youd P, Kennedy A, Gould SR, Darlington LG, Stone TW. Purine, kynurenine, neopterin and lipid peroxidation levels in inflammatory bowel disease. J Biomed Sci. 2002; 9(5):436–42.
63 Yu F, Du Y, Li C, Zhang H, Lai W, Li S, et al. Association between metabolites in tryptophan-kynurenine pathway and inflammatory bowel disease: a two-sample Mendelian randomization. Sci Rep. 2024; 14(1):201.
64 Martin PJ, Storer BE, Levine DM, Hansen JA. Genetic variants associated with inflammatory bowel disease and gut graft-versus-host disease. Blood Advances. 2021; 5(21):4456–64.
65 Zhernakova A, Festen EM, Franke L, Trynka G, van Diemen CC, Monsuur AJ, et al. Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. Am J Hum Genet. 2008; 82(5):1202–10.
66 Lee YH, Song GG. Pathway analysis of a genome-wide association study of ileal Crohn's disease. DNA Cell Biol. 2012; 31(10):1549–54.
67 Ji C, Yang Z, Zhong X, Xia J. The role and mechanism of CARD9 gene polymorphism in diseases. Biomed J. 2021; 44(5):560–6.
68 McGovern DPB, Gardet A, Törkvist L, Goyette P, Essers J, Taylor KD, et al. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet. 2010; 42(4):332–7.
69 Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015; 47(9):979–86.
70 Julià A, Domènech E, Ricart E, Tortosa R, García-Sánchez V, Gisbert JP, et al. A genome-wide association study on a southern European population identifies a new Crohn's disease susceptibility locus at RBX1-EP300. Gut. 2013; 62(10):1440–5.
71 Loddo I, Romano C. Inflammatory Bowel Disease: Genetics, Epigenetics, and Pathogenesis. Front Immunol. 2015; 6: 551.

72 Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. Nat Rev Genet. 2009; 10(1):43–55.
73 Emerson C, Skvarc D, Mikocka-Walus A, Olive L, Gibson PR, Fuller-Tyszkiewicz M. People with Inflammatory Bowel Disease Prefer Cognitive Behavioral Therapy for Fatigue Management: A Conjoint Analysis. Dig Dis Sci. 2024; 69(7):2345–53.
74 Jarmakiewicz-Czaja S, Zielińska M, Sokal A, Filip R. Genetic and Epigenetic Etiology of Inflammatory Bowel Disease: An Update. Genes. 2022; 13(12):2388.
75 Yuan W, Luo Q, Wu N. Investigating the shared genetic basis of inflammatory bowel disease and systemic lupus erythematosus using genetic overlap analysis. BMC Genomics. 2024; 25(1):868.
76 Dal Buono A, Caldirola D, Allocca M. Genetic susceptibility to inflammatory bowel disease: should we be looking to the hypothalamus? Expert Rev Clin Immunol. 2021; 17(8):803–6.
77 Shaw VR, Byun J, Pettit RW, Hou JK, Walsh KM, Han Y, et al. An Atlas Characterizing the Shared Genetic Architecture of Inflammatory Bowel Disease with Clinical and Behavioral Traits. Inflammatory Bowel Diseases. 2024; 30(6):884–93.
78 LeBrun DG, Tran T, Wypij D, Kocher MS. How Often Do Orthopaedic Matched Case-Control Studies Use Matched Methods? A Review of Methodological Quality. Clin Orthop Relat Res. 2019; 477(3):655–62.
79 D Coggon, Geoffrey Rose, DJP Barker. Chapter 8. Case-control and cross sectional studies. In: Epidemiology for the uninitiated. leading general medical journal Research Education Comment. fourth edition. The BMJ; 2020. https://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated/8-case-control-and-cross-sectional
80 Pérez-Guerrero EE, Guillén-Medina MR, Márquez-Sandoval F, Vera-Cruz JM, Gallegos-Arreola MP, Rico-Méndez MA, et al. Methodological and Statistical Considerations for Cross-Sectional, Case-Control, and Cohort Studies. J Clin Med. 2024; 13(14):4005.
81 Mamdani M, Sykora K, Li P, Normand SLT, Streiner DL, Austin PC, et al. Reader's guide to critical appraisal of cohort studies: 2. Assessing potential for confounding. BMJ. 2005; 330(7497):960–2.
82 Khan SM, Tuchman D, Imran A, Lakdawala FM, Mansoor S, Abraham J. A Smooth Transition: Assessing Transition Readiness in Adolescents with Inflammatory Bowel Disease. Dig Dis Sci. 2024;69(10):3640-3649. doi:10.1007/s10620-024-08484-9

Supplementary of Table 1. STROBE ("The Strengthening the Reporting of Observational Studies in Epidemiology") checklist in our observational case-control studies. *Information separately for cases and controls. Abbreviation: NA, Not Applicable; &, and.

	Item No	Recommendation	Page-Line No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	1- 24
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	1- 24 to 33
Introduction			
Background/ rationale	2	Explain the scientific background and rationale for the investigation being reported	2 & 3 - 68 to 98
Objectives	3	State-specific objectives, including any prespecified hypotheses	3- 99 to 109
Methods			
Study design	4	Present key elements of study design early in the paper	3- 111 to 119
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3 & 4- 121 to 153
Participants	6	(*a*) Give the eligibility criteria and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	4 & 5 - 151 to 178 6- 212 to 214
		(*b*) For matched studies, give matching criteria and the number of controls per case	NA
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	5 - 180 to 191
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	3 & 4 – 121 to 137 and 5 - 180 to 210
Bias	9	Describe any efforts to address potential sources of bias	4- 149 to 151 5- 172 to 178
Study size	10	Explain how the study size was arrived at	6 – 212 to 214 & 220 to 222
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	5 - 180 to 210
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	6 - 212 to 230
		(*b*) Describe any methods used to examine subgroups and interactions	5- 170 to 172
		(*c*) Explain how missing data were addressed	6- 213 9- 299 to 300
		(*d*) If applicable, explain how matching of cases and controls was addressed	NA
		(*e*) Describe any sensitivity analyses	NA
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	6- 234 to 244
		(b) Give reasons for non-participation at each stage	NA
		(c) Consider use of a flow diagram	NA

Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	5- 170 to 172
		(b) Indicate number of participants with missing data for each variable of interest	6- 213
Outcome data	15*	Report numbers in each exposure category, or summary measures of exposure	7 to 8- 256 to 273
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	7 to 9- 267 to 305
		(*b*) Report category boundaries when continuous variables were categorized	NA
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	NA
Discussion			
Key results	18	Summarise key results with reference to study objectives	9- 307 to 330
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	12- 456 to 487
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	9 to 12 - 331 to 453 And 13 & 14- 488 to 526
Generalisability	21	Discuss the generalisability (external validity) of the study results	13 & 14- 488 to 526 And 14- 527 to 536
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	NA

Supplementary Table 2. The details regarding the characteristics of the primers for five selected gene variants. Optimal primers have been specifically designed to check the variants of these five selected genes. BLAST approved the design of the primers, and then they were ordered from TAG Copenhagen (www.tagc.com). The details include the number per variant, sequence order, and primer specifications. Various symbols are used to represent different parameters: * (optical absorption at a wavelength of 260 nm), ** (molecular weight calculated based on the oligonucleotide sequence), ǂ (amount of oligonucleotide-based on nanomoles), ϕ (amount of buffer or water required in microliters to make a 100 µM dilution of oligonucleotide),±(oligonucleotide melting temperature), £ (oligonucleotide sequence), and € (number of oligonucleotide bases). All primers were prepared with a concentration of 100 µM, with the degree of synthesis in optical absorption being four times with a molar volume of 0.01 µmol, and desalted purification for all primers. Abbreviation: Outer reverse, OR; Outer forward, OF; Inner Reverse, IR; Inner Forward, Common, C; Mutant, M; Wild type, Wt.

	Primer	OD* (1000µl)	MW** (g/mol)	Nmol ǂ	Water/ tube ϕ (µl/µg)	TM± C°	GC (%)	Seq. (5-3) £	Mer €
1	LACC1-rs3764147IF	5.6	6061	25.6	256/155	49.7	45	GACTCATCATTCCAATGGCA	20
2	LACC1-rs3764147OR	5.2	6743	20.7	207/139	53.0	45.5	AGTATACTTACCAGCGT-GAGCA	22
3	LACC1-rs3764147OF	5.5	7566	20.5	205/155	51.1	32	AAACTCTGCTTTTCTTTG-CATAAAC	25
4	LACC1-rs3764147IR	4.7	6003	23.2	232/139	49.7	45	TTTCTTCCCATAATCCGGAC	20
5	IL17F-rs763780IF	8.1	5974	42.2	422/252	57.9	65	GTCACCCCTGTCATCCGCCA	20
6	IL17F-rs763780OR	8.6	6075	43.4	434/264	53.3	55	GTTTCCATCCGTGCAGGTCT	20
7	IL17F-rs763780OF	8.1	6213	36.2	362/225	53.8	55	AGTGTAGGAACTTGGGCTGC	20
8	IL17F-rs763780IR	8.1	6316	37.4	374/236	50.5	42.9	TATGCACCTCTTACTGCATAC	21
9	GPR35-rs3749171wt	4.9	6087	22.8	228/139	55.9	60	ATGAGCATCAGCCTGGCCAC	20
10	GPR35-rs3749171m	4.0	6102	18.3	183/112	53.8	55	ATGAGCATCAGCCTGGCCAT	20
11	GPR35-rs3749171c	4.8	6138	19.9	199/122	53.8	55	CAGGAAGCAGACCACGAACA	20
12	CARD9-rs4077515M	8.1	6528	31.9	319/208	52.4	47.6	GAACGATGACGAGTGCT-GAAA	21
13	CARD9-rs4077515Wt	8.8	6544	35.1	351/230	54.4	52.4	GAACGATGACGAGTGCT-GAAG	21
14	CARD9-rs4077515C	7.9	6226	37.5	375/233	53.8	55	ACTCTGTGGTTGGGTTTGGG	20
15	GCKR-rs1260326IF	4.2	6093	19.9	199/122	53.8	55	ACCGTGGGTCAGACCTTACT	20
16	GCKR-rs1260326OR	4.3	6053	20.8	208/126	53.8	55	TGGTAACCCATGACCTTGCC	20
17	GCKR-rs1260326OF	5.2	6222	22.6	226/140	53.8	55	AGGGACGGGGTGAATATCCT	20
18	GCKR-rs1260326IR	4.5	6109	21.5	215/131	55.9	60	GCATGGCTGGACTCTCATCG	20

Supplementary Table 3. The gene variant information was utilized to filter and select gene variants for investigation in patient samples. The gene variant chosen was previously linked to IBD disease in research and was associated with amino acid changes. Out of 19 gene variants, 5 were ultimately selected for testing in the patient samples.

ORDER	rs	Our Study (Control)	Azeri	Iranome	World 1000 Genomes	World Gnome AD	World ExAC
1	rs763780 IL17F	0.036	0.055	0.05625	C=0.0935	C=0.0654	C=0.0669
2	rs7530511 IL23R	0.81	0.875	0.8888	C=0.8760	C=0.8422	C=0.8776
3	rs4077515 CARD9	0.226	0.27	0.2819	T=0.3666	T=0.3813	T=0.4019
4	rs3764147 LACC1 (C13orf31)	0.256	0.32	0.2963	G=0.3059	G=0.2673	G=0.2702
5	rs3749171 GPR35	0.0566	0.085	0.1006	T=0.1508	T=0.1902	T=0.1754
6	rs1260326 GCKR	0.526	0.535	0.5869	C=0.7067	C=0.6735	C=0.6428
7	rs2293152 STAT3	0.436	0.6	0.5769	C=0.6538	C=0.6616	-
8	rs1800470 TGFB1	0.496	0.52	0.5088	A=0.5453	-	-
9	rs2228055 IL10RA	0.063	0.07	0.07938	G=0.1144	G=0.0476	G=0.0807
10	rs2075820 NOD1	0.276	0.32	0.3375	T=0.3375	T=0.2802	T=0.2776
11	rs1248696 DLG5	0.89	0.93	0.9368	C=0.9657	C=0.9320	C=0.9311
12	rs721917 SFTPD	0.406	0.48	0.5038	G=0.5046	G=0.4182	G=0.4700
13	rs8108738 MAST3	0.463	0.495	0.4962	A=0.5321	A=0.4703	A=0.5477
14	rs1127354 ITPA	0.06	0.095	0.0993	A=0.0895	A=0.0610	A=0.0770
15	rs738409 PNPLA3	0.243	0.28	0.2525	G=0.2622	gnomAD - Exomes G=0.2778	G=0.2632
16	rs2243639 SFTPD	0.613	0.62	0.6345	C=0.7704	C=0.7117	C=0.6635
17	rs33972313 SLC23A1	0.003	0.015	0.0087	T=0.0353	-	T=0.0293
18	rs2297322 SLC15A1	0.216	0.275	0.26	T=0.3085	T=0.1966	T=0.2137
19	rs34856868 BTBD8	0	0.005	0.0031	A=0.0082	A=0.0200	A=0.0221