The European Research Journal 2024;10(1):8-16 Original Article

DOI: https://doi.org/10.18621/eurj.1285889 *Medical Biology*

Genetically determined plasma trefoil factor-3 levels are causally associated with the risk of ulcerative colitis: a Mendelian randomization study

Bayram Toraman[1], Sami Fidan[2], Gökhan Yıldız[1]

[1]Department of Medical Biology, Karadeniz Technical University Faculty of Medicine, Trabzon, Turkey; [2]Department of Gastroenterology, Karadeniz Technical University Faculty of Medicine, Trabzon, Turkey

ABSTRACT

Objectives: Ulcerative colitis (UC) is an inflammatory disease restricted to the colon's mucosal layer. UC is a complex disease with a largely unknown etiology. Mendelian Randomization (MR) is a method that uses variations in genes that have a causal effect of a modifiable exposure to the disease, in genetic epidemiological studies. Trefoil factor 3 (TFF3) is a secreted protein expressed mainly in the colonic mucosa that binds with the mucin 2 protein, forming a protective barrier for the colon mucosa from bacteria and other insults. This study aimed to identify if TFF3 levels in plasma are causally associated with UC.

Methods: We performed a two-sample MR study. For exposure instrumental variables (IVs), genetically determined TFF3 levels in plasma proteome quantitative trait locus data were obtained from the published literature. Outcome data were obtained from the GWAS catalog. The "TwoSampleMR" R package was used for MR. The statistical significance of IV effect sizes on the outcome is mainly evaluated by the inverse variance weighted (IVW) method.

Results: The IVW test showed considerable statistical significance in all analyzed outcomes except for Crohn's disease (CD) samples. Heterogeneity and horizontal pleiotropy tests showed no significant results for MR sensitivity analysis.

Conclusions: We showed that TFF3 levels in plasma were causally associated with the risk of UC. Increased levels of TFF3 are reversely associated with the risk of UC. The absence of any causal relationship between TFF3 and CD from the same study cohort also supports our causal inference.

Keywords: Ulcerative colitis, trefoil factor-3, genome-wide association study, Mendelian randomization analysis

Inflammatory bowel disease (IBD) is a condition characterized by chronic inflammation of the digestive tract. IBD is separated into two subforms of the condition known as Crohn's disease (CD) and ulcerative colitis (UC) [1]. UC is a usually relapsing auto-inflammatory disease, which is restricted to the mucosal layer of the large intestine [2]. UC is a complex disease, and its etiology is largely unknown. However, several genetic, environmental, and autoimmune factors are suspected to be causative for UC [1]

Corresponding author: Bayram Toraman, PhD., Assistant Professor. Phone: +90 462 377 79 57. E-mail: bayramtoraman@yahoo.com

How to cite this article: Toraman B, Fidan S, Yıldız G. Genetically determined plasma trefoil factor-3 levels are causally associated with the risk of ulcerative colitis: a Mendelian randomization study. Eur Res J. 2024;10(1):8-16. doi: 10.18621/eurj.1285889

Received: April 19, 2023
Accepted: July 21, 2023
Published Online: September 9, 2023

genetic heritable factors play a role in the etiology of IBD [3, 4]. It has also been shown that genetic heritable factors are stronger in CD than in UC [3, 4]. To date, more than 120 genes/genomic loci have been associated with CD [5]. However, although genetics plays a significant role in the cause of UC, only twenty-nine genes, which are cataloged in large databases, were associated with UC [6]. Genetic epidemiological studies use two approaches to identify risk genes or genetic factors in the etiology of complex diseases such as UC. These are hypothesis-driven candidate genes or hypothesis-free genome-wide association studies (GWAS) [7]. GWAS are powerful techniques for identifying risk genes or genetic factors in the etiology of complex traits or diseases [8, 9]. Published GWAS data statistics were deposited in the GWAS catalog (https://www.ebi.ac.uk/gwas/).

Mendelian Randomization (MR) is a method that uses variations in genes that are a causal effect of exposure to disease in genetic epidemiological studies. The MR method has been increasingly used in genetic epidemiological studies in the last decade [10]. MR is a robust technique for confounding factors (sampling bias, environmental confounders, reverse causation, etc.) that cause incomplete and conflicting results in observational epidemiological studies [11]. MR is analog to randomized controlled trials that are the gold standard for clinical causal inference studies. MR technique is based on the second law of Mendel, which is an independent assortment of parental alleles during meiosis and uses genetic variants as instrumental variables (IVs) for causal inference of risk factors onto outcomes (diseases/traits) [10, 11]. MR studies could be performed in a one-sample (single-study samples) or two-sample (two-independent study samples) manner. Two-sample MR (2S-MR) method has more statistical power when considering single-sample MR in terms of reaching a large sample size and needs only summary statistics, which are properly deposited in large databases, such as the GWAS catalog [10, 12].

Trefoil factors are members of the trefoil family of proteins characterized by having at least one copy of the trefoil domain [13]. In the human genome, they are represented by three genes settled in as a cluster on chromosome 21 known as Trefoil factor 1, Trefoil factor 2, and Trefoil factor 3 (TFF3). TFF3 is a secreted protein containing a 40-amino acid trefoil domain connected with three disulfide bonds and expressed mainly in the colonic mucosa (produced by goblet cells) and in other luminal tissues of the breast, thyroid, lung, etc. [13]. In the colonic mucosa, TFF3 binds with mucin 2 protein (MUC2), which is a large glycoprotein, and they form a protective barrier for the colonic mucosa from bacteria and other microorganisms [14]. The functions of TFF3 in IBD, various cancers, and various autoimmune diseases have been extensively studied and the level of its expression is associated with protective or adverse effects with these disorders [15]. In this study, we performed a univariable 2S-MR to find whether genetically determined TFF3 levels were associated with the risk of UC. To our knowledge, this is the first MR study on TFF3 and UC association.

METHODS

Exposure (TFF3) Data (For Discovery Analysis)

Instrumental Variables (Single-Nucleotide Polymorphisms-SNPs) Selections for Exposure

This study hypothesizes that levels of TFF3 in plasma could be associated with the risk of UC disease. Thus, this is a hypothesis-driven gene-centric 2S-MR study. For exposure IVs, genetically determined TFF3 protein levels in plasma proteome quantitative trait locus (pQTL) data (significant SNPs) were obtained from the study of Suhre *et al*. [16]. This study was performed on European individuals (approx. 1,000 individuals) and in this dataset 61 SNPs were associated with plasma TFF3 levels (significance levels ranged from 1.21×10^{-12} to 7.9×10^{-6}). Two of these SNPs were trans (on the other chromosomes) rather than chr6 in which the TFF3 gene is located) and the other SNPs were cis. To obtain independent exposure SNPs from those SNPs located on chr6, we performed a linkage disequilibrium (LD) pair-wise correlation (r^2) matrix by using LDlink (selected the European (CEU) as a reference population) [17]. We pruned out SNPs that had $r^2 \geq 0.30$ on chr6. Then, 6 out of 61 SNPs were determined as independent exposure IVs (Fig. 1, Table 1). F-statistics for each IV were calculated using $F = \frac{R^2 \times (n-2)}{(1-R^2)}$ the formula, where R^2 = explained variance of IVs on exposure, n = sample size of pQTL GWAS exposure [18]. It is generally accepted that the F-statistic of each SNP should be >10 in MR studies

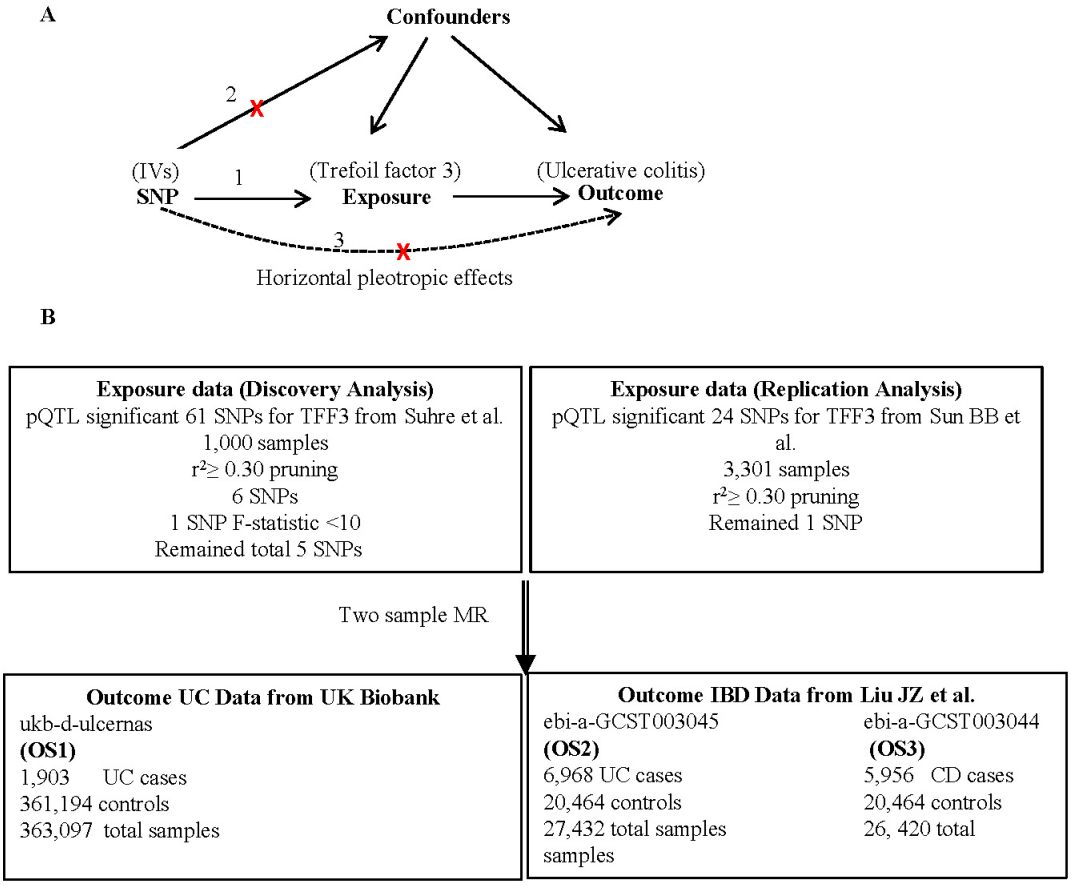

Fig. 1. (A) Mendelian Randomization acyclic graph (causal diagram) 1. SNP is associated with the exposure (relevance), 2. SNP is not associated with confounders (independence), 3. SNP only associated with outcome through the exposure (exclusion restriction). (B) Present study design flowchart.

to avoid weak instrument bias [19]. R^2 was calculated from $2\times EAF\times(1-EAF)\times Beta^2$, where Beta=effect size of IVs, EAF=effect allele frequency of GWAS exposure [18, 20] (Table 1).

Exposure (TFF3) Data (For Replication Analysis)

Independent exposure data for TFF3 plasma levels were obtained from the study of Sun BB *et al*. In this study, 2,994 proteins of plasma in 3,301 individuals from European were analyzed by protein aptamers (SOMAmers) [21]. After pruning only one SNP (rs2524277) remained (Table 1) as an independent SNP and we performed Wald ratio analysis for the effect estimate of MR.

Outcome (UC) Data

For outcome data, we searched the GWAS of UC that the summary statistics deposited in the GWAS catalog, and samples of European descent. The Open GWAS database was especially suitable for this purpose because it has been created for MR studies and contains thousands of summary statistics (https://gwas.mrcieu.ac.uk/). To obtain the most statistical power, we selected two studies that had larger sample sizes and were performed on people of European descent. The first study ID in the Open GWAS database was ukb-d-ulcernas, in which UC GWAS summary-statistic were derived from the United Kingdom (UK) Biobank. These data contained 1,903 UC cases and 361,194 controls (total sample size 363,097). We denoted this study as outcome-study 1 (OS1). The second independent study ID was ebi-a-GCST003045, in which UC GWAS summary statistics were derived from a GWAS meta-analysis study of Liu *et al*. [22] (OS2). OS2 contains 6,968 UC cases and 20,464 population controls of European descent. Additionally, the second study contained 5,956 CD cases (ebi-a-GCST003044) (OS3) and together have

Table 1. SNPs used as instrumental variables in present study

SNP	beta	Se	ea	P value	oa	eaf	R^2	F-statistic
rs2524277	0.7819	0.1086	A	1.22×10^{-12}	G	0.054113	0.0625	66
rs2523586	0.2671	0.0484	G	4.58×10^{-8}	T	0.177316	0.0208	21
rs2523535	0.2205	0.0449	G	1.09×10^{-6}	A	0.249601	0.0182	18
rs3025650	0.4772	0.0987	C	1.55×10^{-6}	T	0.100240	0.0410	42
rs12925077	*-0.3983*	*0.0848*	*A*	*3.08×10^{-6}*	*G*	*0.028155*	*0.0086*	*9*
rs13398473	0.2243	0.0483	T	3.94×10^{-6}	C	0.769369	0.01785	18
rs2524277*	0.4576	0.0506	A	1.58×10^{-19}	G	0.054113	0.02010	70

SNP = single-nucleotide polymorphism, Se = standard error; ea =effect-allele; oa = other allele; eaf = effect-allele frequency, * = Sun *et al.* [21]

been pooled as IBD cases (Fig. 1).

In all original studies in which we used summary statistics of them, it is reported that informed consent and ethical approvals were acquired from the local research ethics committees.

Statistical Analysis

Power was calculated using the total sample size of the outcome data using an online MR-power calculation tool (https://sb452.shinyapps.io/power/) (Table 2). The "TwoSampleMR" (0.5.6), "Mendelian Randomization" (0.5.0), and "'devtools'" (2.4.4) packages were used for MR and statistical analyses. The "ggplot2" package is used for data visualizations. All analyses were conducted on R-platform (version 4.1.0). The statistical significance of IV effect sizes on the outcome is mainly evaluated by the inverse variance weighted (IVW) method. MR-Egger, weighted median, and weighted mode methods were also evaluated as supportive evidence. Heterogeneity and directional (horizontal) pleiotropy of IVs were evaluated using Cochrane's Q statistic and the MR-Egger intercept test, respectively. While reporting our MR study, we have considered the STROBE-MR guidelines (https://www.strobe-mr.org/).

RESULTS

All selected SNPs used in this study are listed in Table 1. These data indicate that rs2524277 has the maximum effect size (beta) and minimum P-value on the TFF3 protein levels in the plasma. The SNP rs12925077 has inverse effect size and the F-statistic of this SNP is 9 (<10), and has a weak-instrument bias; therefore, it has been excluded from subsequent analysis.

IVW test showed considerable statistical significance in all analyzed outcomes except for OS3, which is the CD samples (Table 2). Cochran's Q statistic for testing heterogeneity of IVs by the 2S-MR analysis on OS1 showed no statistically significant heterogeneity (IVW Q=4.00, Q degrees of freedom (Q_df)=4 and Q_pval=0.40; MR Egger Q=3.23, (Q_df)=3 and P= 0.35). Testing for directional horizontal pleiotropy (that is, a violation of the IV assumption 3) by Egger intercept showed no significant horizontal pleiotropy (egger intercept=-0.00018, se=0.00022 and P=0.46). Again, heterogeneity tests of IVW analysis on OS2 and OS3 have not shown any significant heterogeneity (IVW Q=0.03, Q_df=1, Q_pval=0.86 for OS2 andIVW Q= 0.20, Q_df=1, Q_pval=0.65 for OS3). Horizontal pleiotropy could not be evaluated for OS2 and OS3 because only two SNPs remained after the harmonization of SNPs of exposure and outcomes (it should be at least three IVs). The absence of heterogeneity in all study groups also means that the causal relationship is identical between groups (exposure IVs and outcomes). For visual inspection of the results, we created the forest plot and scatter plots from 2S-MR analysis (Fig. 2, Fig. 3, and Fig. 4).

In replication analysis for 2S-MR, we used the plasma proteome data of Sun *et al*. [21]. Wald ratio test showed considerable statistical significance in

Table 2. Results from two sample MR studies performed on three outcomes

Exposure study	Outcome study	Method	nsnp	Beta (se)	OR* (95%CI)	P value	Power
Suhre *et al.* [16] (Discovery)	ukb-d-ulcernas (OS1)	**IVW**	5	-0.0016 (0.0002)	**0.99** (0.98-0.99)	**3.2×10^{-8}**	**0.96**
		Weighted median	5	-0.0014 (0.0004)	0.99 (0.98-1)	**3.7×10^{-4}**	
		MR Egger	5	-0.0011 (0.0006)	0.99 (0.98-1)	0.1850	
		Weighted mode	5	-0.0014 (0.0005)	0.99 (0.98-1)	**0.0327**	
	ebi-a-GCST003045 (OS2)	**IVW**	2	-0.2775 (0.0350)	**0.75** (0.59-0.95)	**2.029×10^{-15}**	**1.00**
		Weighted median	NA#	NA	NA	NA	
		MR Egger	NA	NA	NA	NA	
		Weighted mode	NA	NA	NA	NA	
	ebi-a-GCST003044 (OS3)	**IVW**	2	0.0065 (0.0313)	**1.00** (0.95-1.06)	**0.8333**	**1.00**
		Weighted median	NA	NA	NA	NA	
		MR Egger	NA	NA	NA	NA	
		Weighted mode	NA	NA	NA	NA	
Sun *et al.* [21] (Replication)	ukb-d-ulcernas (OS1)	**Wald ratio**	1	-0.0022 (0.0007)	0.99 (-0.99-1)	**0.0034**	**1.00**
	ebi-a-GCST003045 (OS2)	**Wald ratio**	1	-0.4805 (0.0697)	0.61 (0.53-0.70)	**5.76×10^{-12}**	
	ebi-a-GCST003044 (OS3)	**Wald ratio**	1	0.0258 (0.0623)	1.02 (0.90-1.15)	0.6789	

SNP = single-nucleotide polymorphism, Nsnp = number of SNP, CI = confidence interval, se = standard error, OR = Odds ratio, NA = Not available

*OR was calculated by the formula $OR=e^{\beta}$, where "e" is the base of the natural logarithm, β is the mean effect size of the IVs.

#Because after analysis by the "TwoSampleMR" for OS2 and OS3 there were two SNPs left, outcomes of the weighted median, MR-Egger and weighted mode could not be calculated.

OS2 outcome (OR=0.61; %95 CI 0.53-0.70) except for OS3, which is the CD samples (Table 2). Again, the Wald ratio test showed a statistical significance result in OS1 (OR=0.99; %95 CI-0.99-1). Taken together, we showed that the genetically determined protein levels of the plasma TFF3 are causally associated with the risk of UC in both discovery and replication studies. Increased levels of the plasma TFF3 are inversely correlated with the risk of UC.

DISCUSSION

In this study, we performed 2S-MR to understand whether genetically determined levels of TFF3 in the plasma were associated with the risk of UC. To our knowledge, this is the first MR study on TFF3 and UC association. Also, we could not find any hypothesis-driven observational genetic association study on TFF3 and UC in the published literature. In this study,

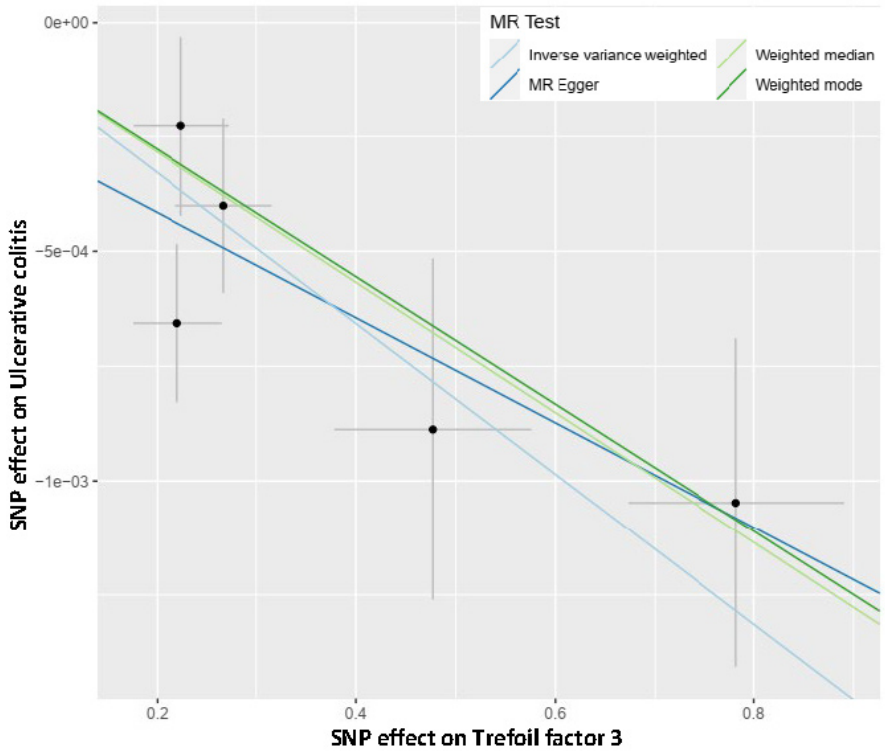

Fig. 2. The scatter plot of causal relationships between Trefoil factor 3 (TFF3) SNPs and ulcerative colitis (UC) in OS1. Each dot represents a SNP, vertical and horizontal black lines around points show 95% confidence intervals (CI) for each polymorphism's associations. The SNPs that increasing TFF3 levels are inversely correlated with the risk of UC. Four different color lines show MR association tests.

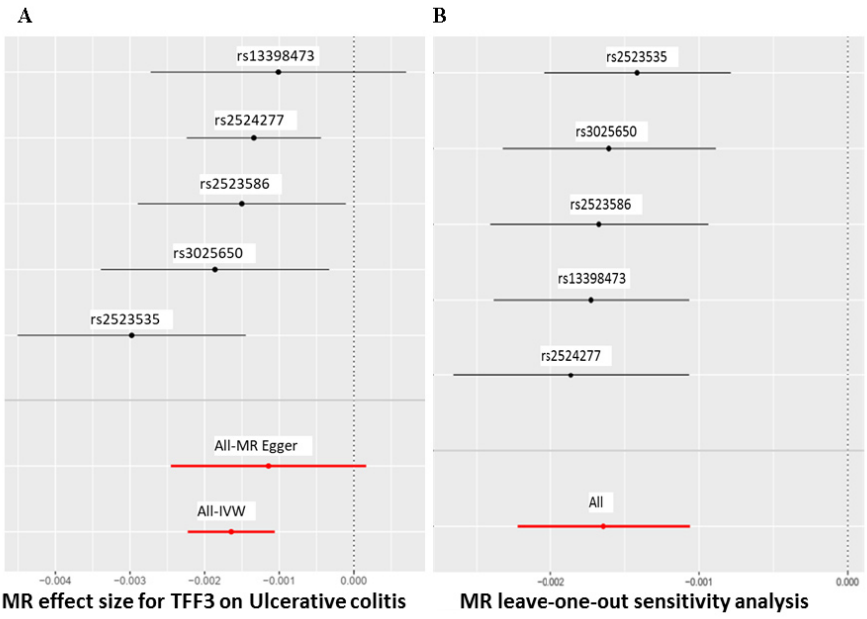

Fig. 3. Forest plots of the OS1 MR results. (A) Forest plot showing the effect estimates of each SNP on UC. (B) Forest plot showing the leave-one-out sensitivity analysis that is the estimated causal effect is shown for each excluded SNP and the overall estimate using all the SNPs. All association is shown in red. Each black dot represents a SNP and horizontal lines around points show 95% CI.

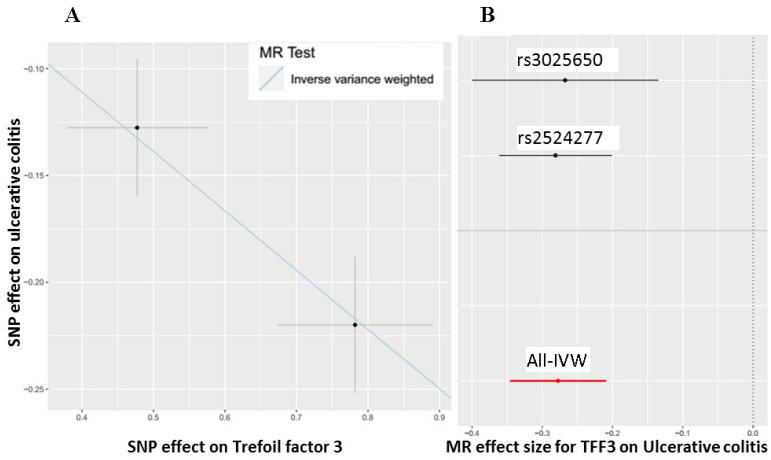

Fig. 4. Scatter and forest plots of the OS2 MR results. (A) The scatter plot of causal relationships between Trefoil factor 3 SNPs and ulcerative colitis. Each dot represents a SNP, vertical and horizontal black lines around points show 95% CI for each polymorphism's associations. (B) Forest plot showing the effect estimates of each SNP on UC. All association is shown in red. Each black dot represents a SNP and horizontal lines around points show 95% CI.

we showed that the effect of increasing TFF3 levels in plasma was associated with the decreasing risk of UC in all study groups of UC (Fig. 2, Fig. 4A). However, the OR of this effect seems quite small in OS1 (OR = 0.99, 95% CI=0.988–0.999) (Table 2). In OS2, the effect of increasing TFF3 levels also decreased the risk of UC (Figure 4A). Nevertheless, the OR was quite larger (significant) than OS1 (OR=0.75, 95% CI=0.59–0.95). One SD increase of the TFF3 level in plasma is associated with a 25 and 39% decreasing the risk of UC in OS2 in the discovery and replication studies, respectively (Table 2). Having more patients with UC in OS2 than OS1 (approximately three-fold) may more accurately reflect the effect of TFF3 on the risk of UC in statistical analysis. Interestingly, in OS3 there was not any effect of TFF3 levels on the risk of CD (OR=1.00, 95% CI=0.95–1.06) (Table 2). These findings support the view that the genetic basis of UC and CD are different [23]. In one study in which it was used 2S-MR approach and performed on the inflammatory proteins in plasma, it was found that an increased level of the vascular endothelial growth factor A (VEGF-A) is associated with a decreased risk of UC [24]. VEGF-A levels were not associated with the risk of CD in that study. In another 2S-MR study by which is performed by Parisinos *et al*. [25], the SNP rs2228145 was used as IV and was showed that the soluble interleukin 6 receptor (sIL6R) levels in plasma are associated with decreased risk of CD and UC. To our knowledge, our 2S-MR study is the 3rd study in which plasma protein levels were used as a biomarker of UC risk assessment. Other studies on plasma TFF3 mainly focused on the possible biomarker role of the plasma TFF3 in the UC disease activity/severity and outcome [26, 27]. Experimental studies performed on the intestinal mucosal barrier function of TFF3 show that TFF3 plays a major role in protecting and repairing the mucosal epithelia from various insults [28]. TFF3 and MUC2 are primarily secreted protein forms of the goblet cells in the colon, they create a protective barrier and a natural immune response in terms of the first line of defense [28]. When Tff3 knockout mice (Tff3-/-) were exposed to dextran sulfate sodium (DSS) induced UC, it has been observed that they had impaired mucosal healing, poor epithelial regeneration, and died from extensive colitis [29]. Thus, our MR study is consistent with the mentioned experimental studies in terms of the crucial role of TFF3 in the etiology and/or pathophysiology of UC. Nevertheless, in the first human clinical study of recombinant TFF3, it was not observed any additional benefits of TFF3 treatment when compared with the effect of corticosteroid treatment alone [30]. However, more clinical trials should be performed to see the exact outcomes of the effects of recombinant TFF3.

Limitations

MR studies are based on some assumptions. Apart

from relevance independence, and exclusion restriction (Fig. 1A), the homogeneity assumption (that is the association of the IVs with the risk factor (TFF3) is homogeneous in the population) is based on acceptance and there is no process of objectively validating it. Our 2S-MR results are based on summary statistics derived from people of European ancestry; it should be careful when extending to other populations due to linkage disequilibrium (LD) differences among different ethnicities. The UC diagnostic criteria and disease definition, as well as other unmeasurable factors, may be different between outcome study groups. Our SNPs, which are used as IVs explain a relatively small variance of the genetically determined plasma TFF3 level (approx. 16%). Therefore, the more independent IVs that explain the greater variance of the levels of TFF3 in plasma will explain the risk of UC more robustly.

CONCLUSION

In the present 2S-MR study, we showed that TFF3 levels in plasma are causally associated with the risk of UC. Increased levels of TFF3 are reversely associated with the risk of UC. The absence of any causal relationship between TFF3 and CD from the same study cohort also supports our causal inference.

Authors' Contribution

Study Conception: BT; Study Design: BT; Supervision: BT; Funding: N/A; Materials: N/A; Data Collection and/or Processing: SF, GY; Statistical Analysis and/or Data Interpretation: BT, SF, GY; Literature Review: BT, SF, GY; Manuscript Preparation: BT and Critical Review: BT, SF, GY.

Conflict of interest

The authors disclosed no conflict of interest during the preparation or publication of this manuscript.

Financing

The authors disclosed that they did not receive any grant during conduction or writing of this study.

REFERENCES

1. Baumgart DC, Carding SR. Inflammatory bowel disease: cause and immunobiology. Lancet. 2007;369(9573):1627-1640. doi: 10.1016/S0140-6736(07)60750-8.
2. Baumgart DC, Sandborn WJ. Inflammatory bowel disease: clinical aspects and established and evolving therapies. Lancet. 2007;369(9573):1641-1657. doi: 10.1016/S0140-6736(07)60751-X.
3. Orholm M, Binder V, Sørensen TI, Rasmussen LP, Kyvik KO. Concordance of inflammatory bowel disease among Danish twins. Results of a nationwide study. Scand J Gastroenterol. 2000;35(10):1075-1081. doi: 10.1080/003655200451207.
4. Tysk C, Lindberg E, Järnerot G, Flodérus-Myrhed B. Ulcerative colitis and Crohn's disease in an unselected population of monozygotic and dizygotic twins. A study of heritability and the influence of smoking. Gut. 1988;29(7):990-996. doi: 10.1136/gut.29.7.990.
5. Garza-Hernandez D, Sepulveda-Villegas M, Garcia-Pelaez J, et al. A systematic review and functional bioinformatics analysis of genes associated with Crohn's disease identify more than 120 related genes. BMC Genomics. 2022;23(1):302. doi: 10.1186/s12864-022-08491-y.
6. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014;42:D1001-1006. doi: 10.1093/nar/gkt1229.
7. Kitsios GD, Zintzaras E. Genome-wide association studies: hypothesis-"free" or "engaged"? Transl Res. 2009;154(4):161-164. doi: 10.1016/j.trsl.2009.07.001.
8. Manolio TA. Genomewide association studies and assessment of the risk of disease. N Engl J Med. 2010;363(2):166-176. doi: 10.1056/NEJMra0905980.
9. Pearson TA, Manolio TA. How to interpret a genome-wide association study. JAMA. 2008;299(11):1335-1344. doi: 10.1001/jama.299.11.1335.
10. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601. doi: 10.1136/bmj.k601.
11. Gagliano Taliun SA, Evans DM. Ten simple rules for conducting a mendelian randomization study. PLoS Comput Biol. 2021;17(8):e1009238. doi: 10.1371/journal.pcbi.1009238.
12. Rasooly D, Patel CJ. Conducting a reproducible mendelian randomization analysis using the R analytic statistical environment. Curr Protoc Hum Genet. 2019;101(1):e82. doi: 10.1002/cphg.82.
13. Gött P, Beck S, Machado JC, Carneiro F, Schmitt H, Blin N. Human trefoil peptides: genomic structure in 21q22.3 and coordinated expression. Eur J Hum Genet. 1996;4(6):308-315. doi: 10.1159/000472224.
14. Järvå MA, Lingford JP, John A, Soler NM, Scott NE, Goddard-Borger ED. Trefoil factors share a lectin activity that defines their role in mucus. Nat Commun. 2020;11(1):2265. doi: 10.1038/s41467-020-16223-7.
15. Yang Y, Lin Z, Lin Q, Bei W, Guo J. Pathological and therapeutic roles of bioactive peptide trefoil factor 3 in diverse diseases: recent progress and perspective. Cell Death Dis. 2022;13(1):62. doi: 10.1038/s41419-022-04504-6.
16. Suhre K, Arnold M, Bhagwat AM, et al. Connecting genetic risk to disease end points through the human blood plasma proteome. Nat Commun. 2017;8:14357. doi: 10.1038/ncomms14357.
17. Machiela MJ, Chanock SJ. LDlink: a web-based application

for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015;31(21):3555-3557. doi: 10.1093/bioinformatics/btv402.
18. Pierce BL, Ahsan H, Vanderweele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. Int J Epidemiol. 2011;40(3):740-752. doi: 10.1093/ije/dyq151.
19. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ. 2018;362:k601. doi: 10.1136/bmj.k601.
20. Ma K, Song P, Liu Z, et al. Genetic evidence suggests that depression increases the risk of erectile dysfunction: A Mendelian randomization study. Front Genet. 2022;13:1026227. doi: 10.3389/fgene.2022.1026227.
21. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558(7708):73-79. doi: 10.1038/s41586-018-0175-2.
22. Liu JZ, van Sommeren S, Huang H, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet. 2015;47(9):979-986. doi: 10.1038/ng.3359.
23. Lawrance IC, Fiocchi C, Chakravarti S. Ulcerative colitis and Crohn's disease: distinctive gene expression profiles and novel susceptibility candidate genes. Hum Mol Genet. 2001;10(5):445-456. doi: 10.1093/hmg/10.5.445.
24. Ek WE, Karlsson T, Höglund J, Rask-Andersen M, Johansson Å. Causal effects of inflammatory protein biomarkers on inflammatory diseases. Sci Adv. 2021;7(50):eabl4359. doi: 10.1126/sciadv.abl4359.
25. Parisinos CA, Serghiou S, Katsoulis M, et al. Variation in interleukin 6 receptor gene associates with risk of Crohn's disease and ulcerative colitis. Gastroenterology. 2018;155(2):303-306.e2. doi: 10.1053/j.gastro.2018.05.022.
26. Nakov R. New markers in ulcerative colitis. Clin Chim Acta. 2019;497:141-146. doi: 10.1016/j.cca.2019.07.033.
27. Nakov R, Velikova T, Nakov V, Ianiro G, Gerova V, Tankova L. Serum trefoil factor 3 predicts disease activity in patients with ulcerative colitis. Eur Rev Med Pharmacol Sci. 2019;23(2):788-794. doi: 10.26355/eurrev_201901_16893.
28. Aamann L, Vestergaard EM, Grønbæk H. Trefoil factors in inflammatory bowel disease. World J Gastroenterol. 2014;20(12):3223-3230. doi: 10.3748/wjg.v20.i12.3223.
29. Mashimo H, Wu DC, Podolsky DK, Fishman MC. Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor. Science. 1996;274(5285):262-265. doi: 10.1126/science.274.5285.262.
30. Mahmood A, Melley L, Fitzgerald AJ, Ghosh S, Playford RJ. Trial of trefoil factor 3 enemas, in combination with oral 5-aminosalicylic acid, for the treatment of mild-to-moderate left-sided ulcerative colitis. Aliment Pharmacol Ther. 2005;21(11):1357-1364. doi: 10.1111/j.1365-2036.2005.02436.x.