Xu *et al. Clinical Proteomics* (2024) 21:59
https://doi.org/10.1186/s12014-024-09511-7

Clinical Proteomics

RESEARCH Open Access

Identification of novel proteins in inflammatory bowel disease based on the gut-brain axis: a multi-omics integrated analysis

Yifeng Xu[1], Zhaoqi Yan[1] and Liangji Liu[2*]

Abstract

Background The gut-brain axis has garnered increasing attention, with observational studies suggesting its involvement in the disease activity and progression of inflammatory bowel disease (IBD), but the precise mechanisms remain unclear.

Materials and methods In this study, we aimed to investigate "novel proteins" underlying IBD in the brain using a comprehensive multi-omics analysis approach. We performed integrated analyses of proteomics and transcriptomics in the human prefrontal cortex (PFC) tissue, coupled with genome-wide association studies (GWAS) of IBD, crohn's disease (CD), and ulcerative colitis (UC). This included performing protein-wide association studies (PWAS), transcriptome-wide association studies (TWAS), Mendelian randomization (MR), and colocalization analysis to identify brain proteins associated with IBD and its subtypes.

Results PWAS analyses identified and confirmation 9, 9, and 6 brain proteins strongly associated with IBD, CD, and UC, respectively. Subsequent MR analyses revealed that increased abundance of GPSM1, AUH, TYK2, SULT1A1, and FDPS, along with corresponding gene expression, led to decreased risk of IBD. For CD, increased abundance of FDPS, SULT1A1, and PDLIM4, along with corresponding gene expression, also decreased CD risk. Regarding UC, only increased abundance of AUH, along with corresponding gene expression, was significantly associated with decreased UC risk. Further TWAS and colocalization analyses at the transcriptome level supported strong associations of SULT1A1 and FDPS proteins with reduced risk of IBD and CD.

Conclusion The two "novel proteins," SULT1A1 and FDPS, are strongly associated with IBD and CD, elucidating their causal relationship in reducing the risk of IBD and CD. This provides new clues for identifying the pathogenesis and potential therapeutic targets for IBD and CD.

Keywords Inflammatory bowel disease, Crohn's disease, Ulcerative colitis, Mendelian randomization, SULT1A1, FDPS

*Correspondence:
Liangji Liu
llj6505@163.com
[1]School of Clinical Medicine, Jiangxi University of Chinese Medicine, Nanchang, Jiangxi, China
[2]Affiliated Hospital of Jiangxi University of Chinese Medicine, Nanchang, Jiangxi, China

Introduction

Inflammatory bowel disease (IBD) represents a chronic inflammation of the gastrointestinal tract, encompassing two main clinical entities: Crohn's disease (CD) and ulcerative colitis (UC). IBD is characterized by a relapsing and remitting course of gastrointestinal symptoms such as rectal bleeding, abdominal pain, and diarrhea [1] The incidence of IBD is steadily increasing in regions like North America, Oceania, and Europe [2]. However, despite ongoing research efforts, the exact etiology of IBD remains incompletely understood, and there is currently no curative therapy available. Additionally, issues such as medication resistance and adverse drug reactions underscore the urgent need for novel therapeutic approaches [3].

In recent years, the intricate and extensive relationship between IBD and the nervous system has garnered increasing attention, leading to the conceptualization of the gut-brain axis (GBA). Evidence suggests a bidirectional interaction between GBA and IBD [4, 5], with meta-analyses supporting these findings [6]. Numerous studies have implicated psychological stress in modulating intestinal permeability, motility, sensitivity, gut microbiota composition, as well as promoting the development and reactivation of intestinal inflammation [7–9], ultimately influencing the course of IBD [10]. Intestinal inflammation is also believed to alter brain activity and behavior through GBA regulation, potentially triggering psychiatric and neurodegenerative disorders [11–13]. Consequently, there is increasing interest in understanding the impact of the nervous system on IBD pathogenesis and exploring IBD therapeutic strategies based on the GBA concept [14].

Genetic susceptibility is a hallmark of IBD, with early genetic studies indicating a relative risk of over fivefold for first-degree relatives of IBD patients compared to the general population [15]. Significant advancements in IBD genetics have been achieved with the rapid development of genomic technologies such as whole-exome sequencing and genome-wide association studies (GWAS), identifying over 240 risk loci that shed light on various pathogenic mechanisms involving bacterial recognition, inflammation, immunity, and autophagy [16]. However, our understanding of the genetic and proteomic associations with IBD in the central nervous system remains limited, hindered by challenges in conducting relevant animal experiments. Integration of GWAS data with multidimensional quantitative trait locus (QTL) data holds promise for identifying specific genes or proteins underlying IBD pathogenesis [17].Previous research has explored therapeutic targets for inflammatory bowel disease using drug-target Mendelian randomization (MR) [18], while other scholars have utilized Summary data-based MR (SMR) analysis to investigate therapeutic targets related to important pathological mechanisms like endoplasmic reticulum stress in UC and CD [19]. However, research on the genetic and proteomic associations with IBD in the central nervous system based on the gut-brain axis remains limited.

In this study, we aim to investigate novel proteins underlying IBD in the brain using a comprehensive multi-omics approach. Firstly, we integrate two sets of protein quantitative trait locus (pQTL) data obtained from human brain tissue and summary statistics from IBD GWAS to conduct a proteome-wide association study (PWAS) to identify proteins associated with IBD. Subsequently, we employ two-sample MR analysis to confirm causal relationships between significant proteins and their upstream mRNA and utilize genetic colocalization analysis to assess the probability of two traits sharing the same causal variants. Finally, we validate the association of key proteins at the mRNA level through transcriptome-wide association analysis (TWAS). Proteins passing all these tests will be identified as novel players in the pathogenesis of IBD (Fig. 1).

Materials and methods

Data sources

Human brain proteomics data

For this study, a total of four sets of QTL datasets were utilized as genetic instrumental variables, consisting of two pQTL and two eQTL datasets. Both pQTL datasets underwent the same quality control procedures to identify and control for clinical covariates (i.e., age, sex, and final clinical diagnosis of cognitive status) before estimating protein weights. Further details regarding sample description, proteomic analysis, quality control, and statistical analysis are available in the original publications.

pQTL in the discovery PWAS: The discovery dataset was obtained from Wingo et al.'s study conducted at the Banner Sun Health Research Institute (Banner) [20], comprising 198 individuals of European ancestry diagnosed with either Alzheimer's disease (AD) or normal cognition. Following quality control, a subset of 152 individuals with dorsal lateral prefrontal cortex (dlPFC) tissue samples was selected for quantitative assessment of 11,518 proteins, with roughly balanced gender distribution (87 females, 65 males), and an average age at death of 85 years.

pQTL in the confirmation PWAS: The validation dataset ROSMAP, also developed by Wingo et al. [21], originated from the Religious Orders Study and Rush Memory and Aging Project (ROS/MAP) cohorts, recruiting 391 individuals from two longitudinal clinical-pathological studies on aging and AD. After quality control, the analysis included dlPFC tissue samples from 376 individuals of European ancestry.

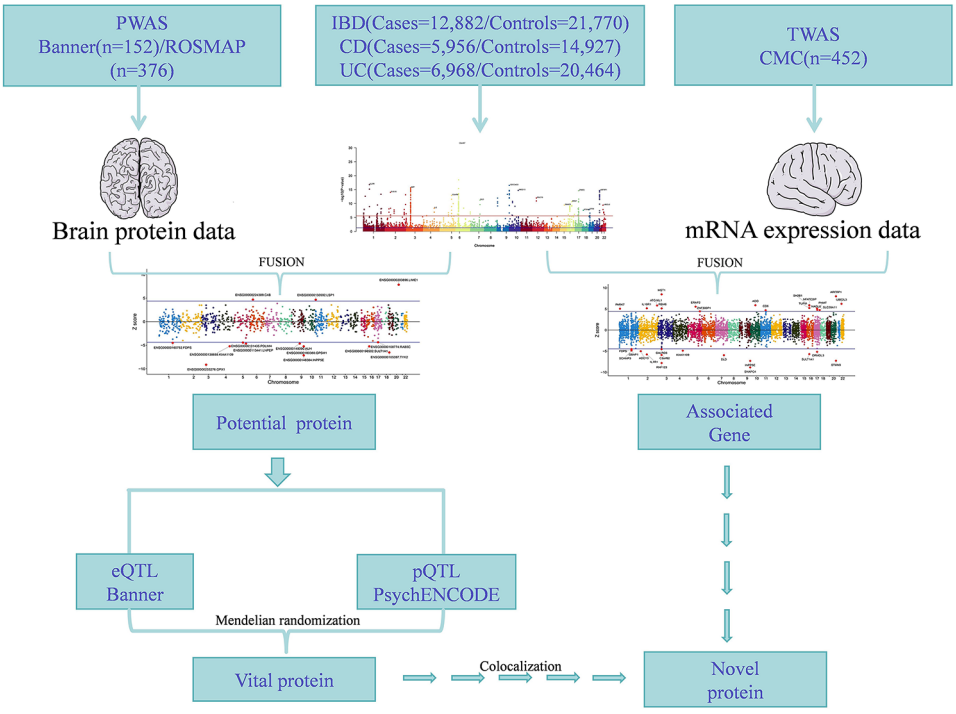

Fig. 1 Research flow chart. IBD: Inflammatory Bowel Disease; CD: Crohn's Disease; UC: Ulcerative Colitis; PWAS: Proteome-wide association studies; TWAS: Transcriptome-wide association studies; pQTL: protein Quantitative Trait Loci; eQTL: expression Quantitative Trait Loci

Human brain transcriptomics data

The eQTL dataset from PsychENCODE was generated from the prefrontal cortex (PFC) of 1,378 human individuals [22], incorporating 2,542,908 single nucleotide polymorphism(SNP)-gene expression pairs adjusted for 50 PEER factors. Summarized data can be downloaded from the SMR website (https://cnsgenomics.com/software/smr/#eQTLsummarydata).

eQTL in the verification TWAS: The transcriptomic data for TWAS in this study were sourced from the CommonMind Consortium (CMC), utilizing postmortem dlPFC tissue samples from 452 individuals of European ancestry. Genotyping of samples was performed using the Illumina Infinium HumanOmniExpress Exome array, and gene expression was mapped to human Ensembl genes using TopHat v.2.0.9 and Bowtie v.2.1.0 after phasing and imputation. We selected the cis-eQTL dataset, and detailed information regarding quality control, data adjustment, and normalization procedures can be found in the study by Fromer et al. [23]. The summarized statistical data above is from publicly available websites and does not include any personal identifying information of participants, therefore, no additional ethical approval is required for this study.

GWAS summary data

We included GWAS summary data for IBD and its two major subtypes, sourced from the International IBD Genetics Consortium (IIBDGC) [24], comprising 15 European ancestry cohorts. The overall IBD dataset included 12,882 cases and 21,770 controls, while CD comprised 5,956 cases and 14,927 controls, and UC comprised 6,968 cases and 20,464 controls. Cases included in the IIBDGC were diagnosed using standard clinical, endoscopic, and histopathological criteria. The summary statistics data for GWAS is sourced from publicly available websites and does not contain any personal identifying information of participants. Therefore, no additional ethical approval is required for this study.

Statistical analysis

The PWAS and TWAS analyses were conducted by integrating GWAS data for IBD and its two subtypes with two sets of brain proteomics data (discovery and confirmation) and one set of transcriptomic data obtained from brain tissue using the FUSION pipeline (http://gusevlab.org/projects/fusion/). We conducted PWAS analysis to discover and validate brain proteomes associated with IBD. Given that proteins are translated from mRNA, we also explored genes associated with IBD at the transcriptome (mRNA) level by TWAS [25]. Specifically, we first calculated weights for the proteome and transcriptome separately using FUSION. Subsequently, PWAS and TWAS analyses were performed by computing the linear sum of independent SNP z-scores multiplied by weights to combine the genetic effects of IBD and its subtypes (GWAS z-scores) with pre-computed weights. Proteins identified in the discovery dataset and replicated

in the confirmation dataset were considered "potential proteins" associated with IBD. To ensure the reliability of FUSION mapping and control for potential effects of multiple testing on PWAS results, we implemented over 10,000 permutations to correct for population stratification and inflation or deflation of the empirical null distribution in GWAS, applying the false discovery rate (FDR) using the Benjamini-Hochberg (BH) method with permutation proteins with FDR-adjusted $P<0.05$ were considered significant. Conditional quantile–quantile (Q–Q) plots were used to visualize FDR P-values to assess enrichment of pleiotropy and shared risk loci for genes associated with IBD. Additionally, due to structural complexity, the major histocompatibility complex (MHC) region on chromosome 6 was excluded from the analysis. The linkage disequilibrium reference panel commonly used for this study comprised 1,190,321 HapMap SNPs from 489 individuals of European ancestry in the 1000 Genomes Project.

Mendelian randomization analysis

MR is a genetic method that investigates causal relationships between exposure and outcome using SNPs significantly associated with exposure, closely correlated with genetic instrumental variables (IVs) across the entire genome. This study adheres to the STROBE-MR guidelines [26]. In this study, IVs for exposure were derived from the eqtl and pqtl datasets, representing proteins "associated protein" with significant associations with IBD and its subtypes tested through double verification in PWAS (discovery and confirmation). We selected SNPs with genome-wide significance ($P<5\times10^{-8}$) for these genes and aggregated them within 10 MB distance with LD $R^2\leq0.001$ as IVs. Next, IVs were extracted from GWAS summary data for outcome traits (IBD and its two subtypes) and harmonized in exposure and outcome GWAS datasets. Generally, MR effects were evaluated using the inverse variance-weighted (IVW) method. However, after clumping, most pQTLs or eQTLs had only one SNP, at which point the "Wald ratio" method was used for MR assessment. Proteins were deemed "vital protein" if both eQTLs and pQTLs were causally associated with IBD and its subtypes. The aforementioned algorithm was implemented using the "Twosample MR" R package.

Colocalization analysis

We further conducted colocalization analysis using the colco.abf function in the R package "Coloc" to assess the likelihood of shared causal signals between risk loci for IBD and its subgroups and pQTLs or eQTLs. "Coloc" employs Bayesian algorithms to generate posterior probabilities for five mutually independent variables, where posterior probability PPH3+PPH4>0.8 is typically interpreted as colocalization [27]. The combined significance from PWAS, MR, colocalization, and TWAS analyses greatly ensured the accuracy of the impact of "vital protein" on IBD and ultimately identified "novel protein".

Results

Discovery and confirmation PWAS for IBD

In the discovery phase of PWAS (Banner), we identified 14 proteins whose levels in the brain were significantly associated with ($P\leq0.05/1147=4.36E-05$) susceptibility to IBD. Among these 14 proteins, the minimum absolute Z-value was 4.51, and the minimum P-value was 6.63E-06. Subsequently, independent confirmatory PWAS using a different reference brain proteome (ROSMAP) replicated 9 genes ($P\leq0.05/1761=2.84E-05$), accounting for a high proportion of 64.3% (Table 1; Fig. 2A; Additional file1 Table S1).

For CD, in the discovery PWAS (Banner), we found 14 proteins whose levels were associated with susceptibility to CD ($P\leq0.05/1147=4.36E-05$). Among these 14 proteins, the minimum absolute Z-value was 4.13, and the minimum P-value was 3.69E-05. Subsequently, independent confirmatory PWAS replicated 9 genes ($P\leq0.05/1761=2.84E-05$), accounting for a high proportion of 64.3% (Table 1; Fig. 2B; Additional file1 Table S2).

In the case of UC, in the discovery PWAS (Banner), we identified 9 proteins whose levels were associated with susceptibility to UC ($P\leq0.05/1147=4.36E-05$). Among these 9 proteins, the minimum absolute Z-value was 4.09, and the minimum P-value was 4.23E-05. Subsequently, independent confirmatory PWAS replicated 6 genes ($P\leq0.05/1761=2.84E-05$), accounting for a high proportion of 66.7% (Table 1; Fig. 2C; Additional file1 Table S3).

Finally, consistent patterns were observed in all discovery and confirmation conditional Q-Q plots, showing significant upward deviation, indicating significant pleiotropic enrichment of dlPFC brain tissue proteins with IBD and related phenotypes (Fig. 3).

MR analysis

Two-sample MR further confirms the causal relationship between the "Banner dataset" and "potential proteins" in brain proteomics and IBD and its subtypes. Using the "Wald ratio" method, 9 "associated proteins" were found to be causally associated with IBD, with most proteins reducing the risk of IBD. Specifically, an increase in abundance of INPP5E, GPSM1, AUH, LNPEP, GPX1, TYK2, SULT1A1, and FDPS led to a decrease in IBD risk, while only an increase in LSP1 abundance resulted in an increased risk of IBD. In CD, LIME1, LNPEP and MSTO1 did not have genome-wide significant SNPs ($P<5\times10^{-8}$) and were therefore not included in MR analysis. However, the remaining 6 proteins were all causally associated with CD, with significant associations observed

Table 1 Results of PWAS and co-localization analysis identifying Brain proteins Associated with IBD and its subtypes

Information				Discovery PWAS(Banner)				Confirmatory PWAS(ROSMAP)			
Protein	**CHR**	**P0**	**P1**	**PWAS_Z**	**PWAS_P**	**PP3**	**PP4**	**PWAS_Z**	**PWAS_P**	**PP3**	**PP4**
IBD											
AUH	9	93,976,097	94,124,195	-4.64	3.45E-06	0.01	0.36	-4.79	1.68E-06	0.01	0.97
FDPS	1	155,278,539	155,290,457	-4.51	6.63E-06	0.04	0.93	-4.51	6.63E-06	0.04	0.93
GPSM1	9	139,221,932	139,254,057	-5.24	1.62E-07	0.03	0.08	-4.58	4.56E-06	1.00	0
GPX1	3	49,394,609	49,396,033	-9.14	6.50E-20	0.01	0.99	-8.83	1.09E-18	0.01	0.99
INPP5E	9	139,323,071	139,334,274	-7.15	8.91E-13	0.20	0.18	-7.03	2.11E-12	0.99	0
LIME1	20	62,366,815	62,370,456	7.84	4.34E-15	0.01	0.56	7.87	3.68E-15	0.01	0.77
LNPEP	5	96,271,098	96,373,219	-4.44	9.08E-06	0.05	0.93	-4.94	7.96E-07	0.04	0.96
LSP1	11	1,874,200	1,913,497	4.65	3.31E-06	2.00E-03	0.55	4.46	8.30E-06	2.00E-03	0.96
SULT1A1	16	28,616,903	28,634,946	-5.24	1.59E-07	0.12	0.89	-4.66	3.10E-06	0.12	0.88
CD											
FDPS	1	155,278,539	155,290,457	-5.43	5.72E-08	0.02	0.98	-5.43	5.72E-08	0.02	0.98
GPX1	3	49,394,609	49,396,033	-6.85	7.14E-12	0.05	0.95	-6.28	3.44E-10	0.03	0.97
HINT1	5	130,494,720	130,507,428	-4.40	1.09E-05	0.04	0.94	-4.70	2.57E-06	0.04	0.93
INPP5E	9	139,323,071	139,334,274	-6.18	6.34E-10	0.20	0.17	-6.26	3.80E-10	0.99	0
LIME1	20	62,366,815	62,370,456	6.19	6.01E-10	0.02	0.39	6.28	3.40E-10	0.02	0.71
LNPEP	5	96,271,098	96,373,219	-6.01	1.88E-09	0.09	0.89	-5.69	1.24E-08	0.03	0.97
MSTO1	1	155,579,979	155,718,153	-4.82	1.40E-06	0.01	0.97	-4.77	1.86E-06	0.04	0.70
PDLIM4	5	131,593,364	131,609,147	-4.84	1.28E-06	1.00	0.00	-4.30	1.72E-05	1.00	0.00
SULT1A1	16	28,616,903	28,634,946	-5.11	3.30E-07	0.08	0.92	-4.51	6.39E-06	0.11	0.89
UC											
AUH	9	93,976,097	94,124,195	-4.13	3.61E-05	2.00E-03	0.22	-4.22	2.48E-05	0.01	0.95
GPX1	3	49,394,609	49,396,033	-7.00	2.48E-12	0.01	0.99	-7.03	2.01E-12	0.01	0.99
INPP5E	9	139,323,071	139,334,274	-4.60	4.29E-06	0.20	0.18	-4.39	1.11E-05	0.99	0
LIME1	20	62,366,815	62,370,456	5.51	3.60E-08	0.01	0.79	5.25	1.50E-07	0.02	0.68
LSP1	11	1,874,200	1,913,497	5.01	5.46E-07	3.00E-03	0.93	4.93	8.43E-07	2.00E-03	1.00
PANK4	5	131,593,364	131,609,147	-4.40	1.09E-05	0.87	0.14	-4.74	2.16E-06	0.97	0.03

CHR: Chromosome; P0: Start point in base pairs; P1: End point in base pairs; PP, Proportional p-value; IBD: Inflammatory Bowel Disease; CD: Crohn's Disease; UC: Ulcerative Colitis

for an increase in abundance of GPX1, INPP5E, FDPS, SULT1A1, PDLIM4, and HINT1, resulting in a decrease in CD risk. As for UC, LIME1 did not have genome-wide significant SNPs, but the remaining 5 proteins were causally associated with IBD. Among these, an increase in abundance of GPX1, INPP5E, PANK4, and AUH was associated with a decrease in UC risk, while only an increase in the abundance of LSP1 is associated with an increased risk of UC (Table 2).

The PsychENCODE eQTL summary dataset was used to further examine whether these "associated proteins" remained causally associated with IBD at the gene (mRNA) level. Specifically, we validated transcriptomic data for 7, 3, and 3 proteins in IBD, CD, and UC, respectively. For IBD, the LSP1 gene did not show a significant causal relationship with IBD, while the AUH, TYK2, SULT1A1, FDPS, and GPSM1 genes exhibited the same direction of causal relationship as the expressed brain tissue proteins. In CD, 3 genes, FDPS, SULT1A1, and PDLIM4, showed the same direction of causal relationship as the expressed brain tissue proteins. For UC, only the AUH gene showed the same direction of causal relationship as the expressed brain tissue proteins, while the LSP1 and PANK4 genes did not show a significant causal relationship with UC (Table 3).

Colocalization and TWAS validation

Colocalization analysis was used to scan for shared pathogenic genes between IBD, CD, and UC in the above conclusions. The results showed that most of the proteins confirmed by PWAS shared causal variant drivers with IBD and its subgroups, with FDPS, GPX1, LNPEP, and SULT1A1 shared in IBD, FDPS, GPX1, HINT1, LNPEP, PDLIM4, and SULT1A1 shared in CD, and GPX1, LSP1, and PANK4 shared in UC. Finally, TWAS provided secondary validation for the "vital protein" SULT1A1 and FDPS, both strongly associated with IBD and CD, while other genes did not yield significant findings (Table 4; Fig. 4, Additional file1 Table S4). Notably, the genes SULT1A1 and FDPS, along with their expressed proteins, shared common causal variant drivers with IBD and CD, thus ultimately identified as "novel proteins".

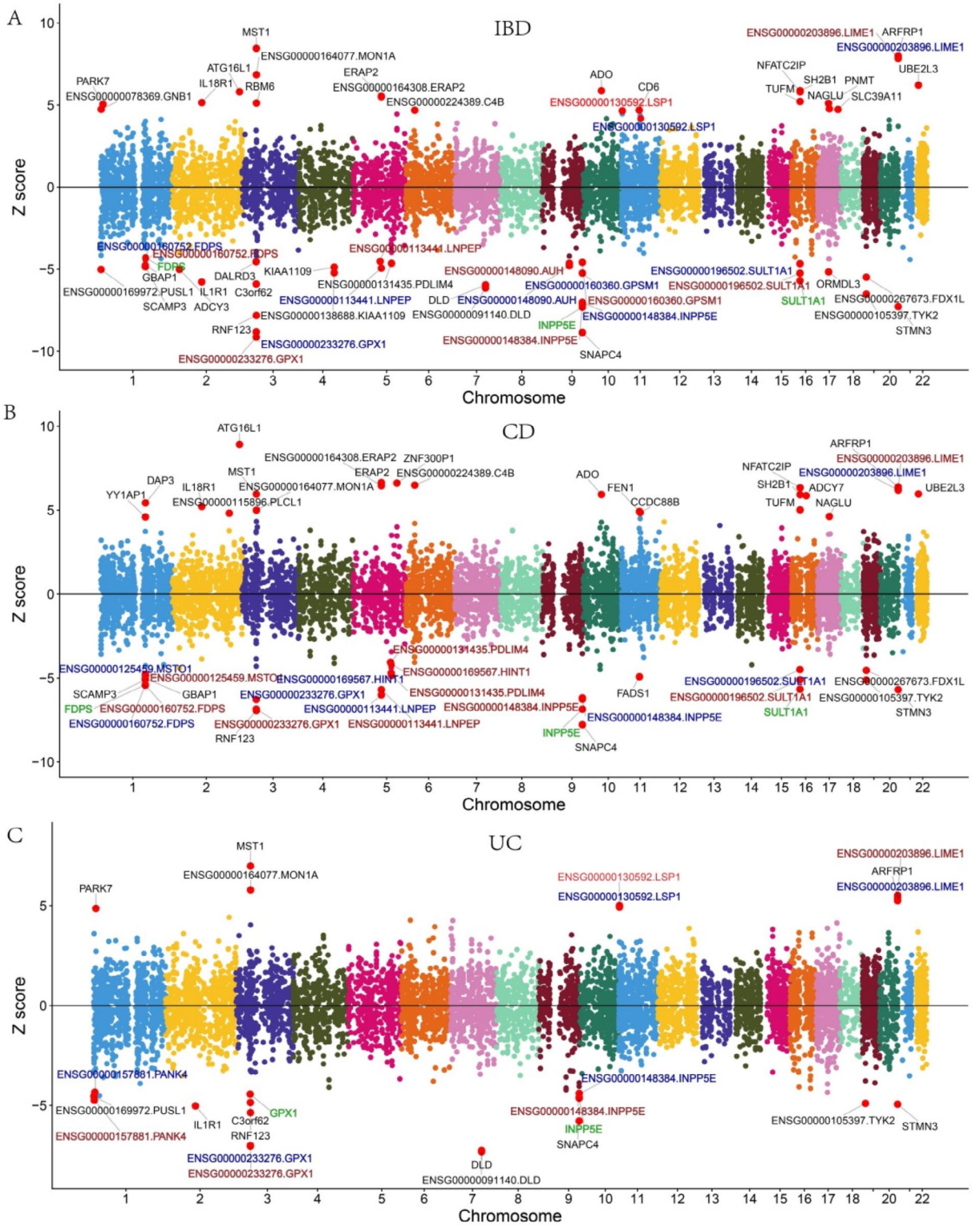

Fig. 2 Manhattan plots of PWAS and TWAS for inflammatory bowel disease and its subtypes. Manhattan plots depict significant human brain proteins and genes discovered in PWAS and TWAS for IBD and its subtypes. Each point represents a single test of association between a gene and phenotype, plotted based on genomic position on the x-axis and Z-value on the y-axis. Statistically significant proteins and genes identified through analysis are plotted, with annotated regions denoting those that have been validated. Proteins identified in both discovery PWAS and confirmation PWAS are highlighted in red and blue fonts, with red representing Banner in discovery PWAS and blue representing ROSMAP in confirmation PWAS. Green font indicates genes significantly identified in TWAS analysis of proteins validated by PWAS. Subplots are labeled for IBD, CD, and UC

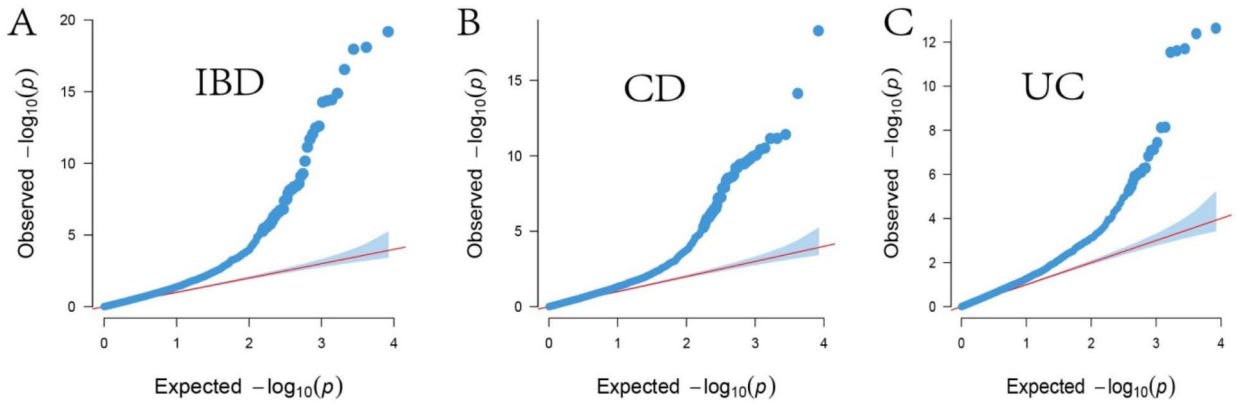

Fig. 3 Conditional quantile-quantile plots. The dashed line represents the expected line under the null hypothesis, with leftward deviation indicating the degree of pleiotropic enrichment. (**A**) Conditional Q-Q plot summarizing PWAS/TWAS and IBD. (**B**) Conditional Q-Q plot summarizing PWAS/TWAS and CD. (**C**) Conditional Q-Q plot summarizing PWAS/TWAS and UC. IBD: Inflammatory Bowel Disease; CD: Crohn's Disease; UC: Ulcerative Colitis

Table 2 Mendelian randomization analysis validating causal relationships at the proteomic level

Exposure	Method	Beta	Se	Pval	OR(95%CI)
IBD					
LSP1	Wald ratio	0.44	0.10	8.30E-06	1.55(1.28,1.89)
LNPEP	Wald ratio	-2.20	0.47	3.54E-06	0.11(0.04,0.28)
TYK2	Wald ratio	-0.93	0.14	6.75E-11	0.39(0.29,0.52)
GPX1	Wald ratio	-1.35	0.15	6.50E-20	0.25(0.19,0.34)
SULT1A1	Wald ratio	-0.38	0.07	9.15E-09	0.68(0.60,0.77)
FDPS	Wald ratio	-0.81	0.18	6.63E-06	0.44(0.31,0.63)
GPSM1	Wald ratio	-1.48	0.47	1.56E-03	0.22(0.09,0.56)
AUH	Wald ratio	-1.57	0.34	3.45E-06	0.20(0.10,0.40)
INPP5E	Wald ratio	-1.84	0.26	2.11E-12	0.15(0.09,0.26)
CD					
FDPS	Wald ratio	-1.32	0.24	5.72E-08	0.26(0.16,0.42)
INPP5E	Wald ratio	-2.22	0.35	3.80E-10	0.10(0.05,0.21)
PDLIM4	Wald ratio	-0.72	0.18	4.87E-05	0.48(0.34,0.68)
GPX1	Wald ratio	-1.38	0.20	7.14E-12	0.25(0.17,0.37)
SULT1A1	Wald ratio	-0.50	0.09	3.37E-08	0.60(0.51,0.72)
HINT1	Wald ratio	-2.19	0.48	6.16E-06	0.11(0.04,0.28)
UC					
GPX1	Wald ratio	-1.30	0.19	2.48E-12	0.27(0.19,0.39)
LSP1	Wald ratio	0.59	0.12	8.43E-07	1.80(1.42,2.28)
AUH	Wald ratio	-1.76	0.43	3.61E-05	0.17(0.07,0.39)
PANK4	Wald ratio	-1.70	0.73	0.02	0.18(0.04,0.76)
INPP5E	Wald ratio	-1.44	0.33	1.11E-05	0.23(0.12,0.45)

Mendelian Randomization analysis was conducted to validate the causal relationships between IBD, CD, UC, and the brain proteins identified by PWAS

Table 3 Mendelian randomization analysis validating causal relationships at the transcriptomic level

Exposure	Method	Beta	Se	Pval	OR(95%CI)
IBD					
AUH	Wald ratio	-0.15	0.04	9.00E-04	0.86(0.79,0.94)
GPSM1	Wald ratio	-0.35	0.06	3.97E-08	0.70(0.62,0.79)
FDPS	Wald ratio	-0.23	0.05	2.90E-06	0.79(0.72,0.87)
SULT1A1	Wald ratio	-0.16	0.03	1.31E-08	0.85(0.81,0.90)
LNPEP	Wald ratio	0.16	0.07	0.02	1.17(1.03,1.34)
LSP1	Wald ratio	-0.05	0.04	0.28	0.95(0.88,1.03)
PDLIM4	Wald ratio	-0.09	0.02	4.51E-05	0.91(0.87,0.95)
TYK2	Wald ratio	-0.25	0.07	3.32E-04	0.78(0.68,0.89)
CD					
FDPS	Wald ratio	-0.37	0.07	2.08E-08	0.68(0.60,0.78)
PDLIM4	Wald ratio	-0.13	0.03	1.13E-05	0.87(0.82,0.92)
SULT1A1	Wald ratio	-0.21	0.04	9.80E-09	0.80(0.75,0.86)
ADO	Wald ratio	0.48	0.07	1.91E-13	1.61(1.42,1.84)
TYK2	Wald ratio	-0.39	0.09	3.87E-05	0.68(0.56,0.81)
UC					
KIAA1109	Wald ratio	-0.20	0.05	1.34E-04	0.81(0.74,0.90)
AUH	Wald ratio	-0.17	0.06	3.02E-03	0.84(0.75,0.94)
PANK4	Wald ratio	0.15	0.09	0.10	1.16(0.97,1.39)
LSP1	Wald ratio	0.00	0.05	0.95	1.00(0.90,1.11)

Mendelian Randomization analysis was conducted to validate the causal relationship between IBD, CD, UC, and the transcriptomic levels of brain proteins identified by TWAS

Discussion

This study represents the first comprehensive exploration of the genetic correlations between IBD and its two subtypes with the brain proteins. The current study has several strengths. Firstly, the design of discovery and validation based on two sets of proteomic data from the PFC enhances the robustness of our investigative findings, preliminarily identifying "potential proteins" associated with IBD and its subtypes. These findings were then validated at the protein and mRNA levels through MR, assessing their causal associations, followed by colocalization analysis evaluating causal variant drivers. Finally, TWAS based on mRNA data from the PFC was utilized to further validate genes associated with IBD and its subtypes among these candidate proteins. Our results indicate that 9 "associated proteins" and their "associated genes" are causally associated with IBD, with 6 causally associated with CD and 5 with UC. Colocalization analysis suggests the presence of common causal variants

Table 4 Transcriptomic levels of brain proteins associated with inflammatory bowel disease and its subtypes in TWAS-Based secondary validation PWAS analysis

Gene	CHR	P0	P1	TWAS_Z	TWAS_P	PP3	PP4
IBD							
INPP5E	9	139,323,071	139,334,274	-7.31719	2.53E-13	1	0
SULT1A1	16	28,616,903	28,634,946	-5.69911	1.20E-08	0.197	0.803
FDPS	1	155,278,539	155,290,457	-4.5129	6.40E-06	0.043	0.796
CD							
INPP5E	9	139,323,071	139,334,274	-6.8575	7.01E-12	1.00	0
SULT1A1	16	28,616,903	28,634,946	-5.66809	1.44E-08	0.16	0.844
FDPS	1	155,278,539	155,290,457	-5.400534	6.64E-08	0.02	0.934
UC							
GPX1	3	49,394,608	49,395,791	-4.4621	8.12E-06	1	0
INPP5E	9	139,323,066	139,334,256	-4.632611	3.61E-06	1	0

CHR: Chromosome; P0: Start point in base pairs; P1: End point in base pairs; PP, Proportional p-value; IBD: Inflammatory Bowel Disease; CD: Crohn's Disease; UC: Ulcerative Colitis

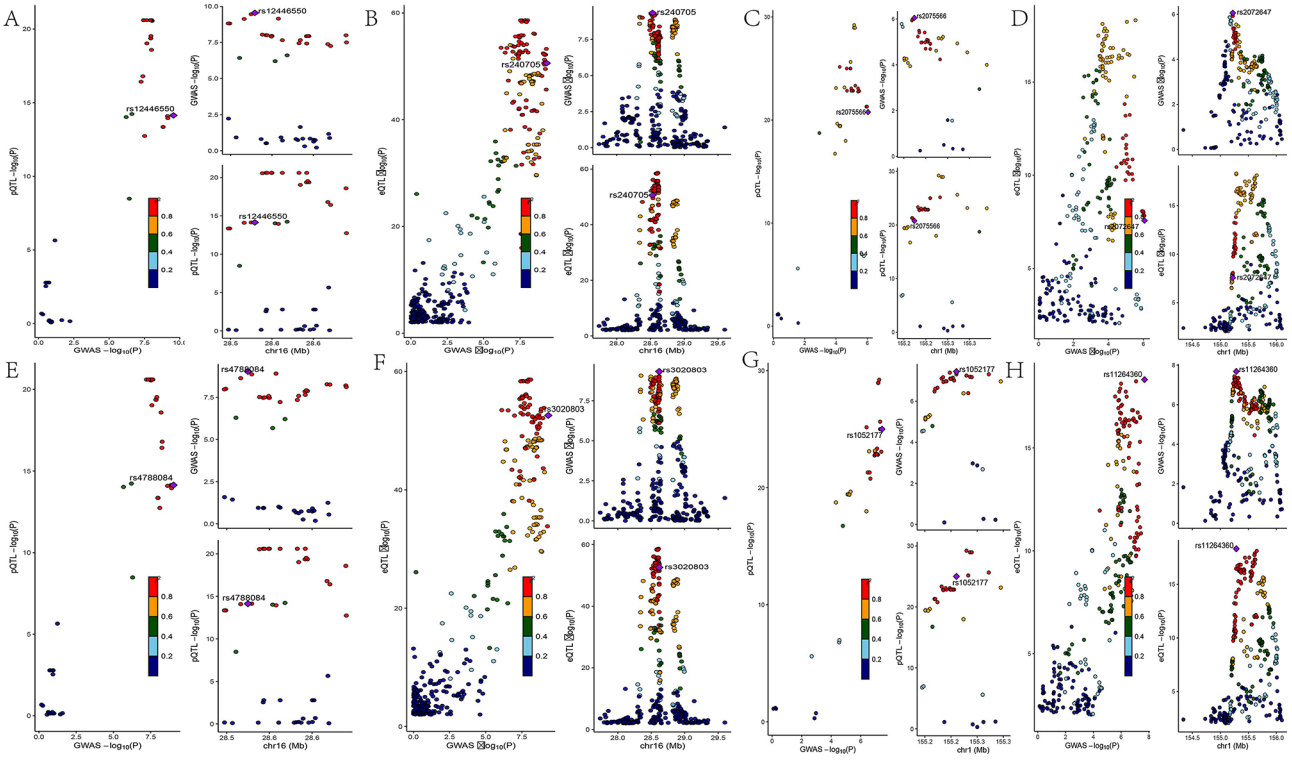

Fig. 4 Co-localization analysis plot. Illustration of co-localization results validated through PWAS, MR, and TWAS in IBD and CD. The upper and lower portions of the figure represent sections for IBD and CD, respectively. Annotations denote: (**A**) IBD-associated SULT1A1 protein from discovery pQTLs data; (**B**) IBD-associated SULT1A1 gene from PsychENCODE eQTLs data; (**C**) IBD-associated FDPS protein from discovery pQTLs data; (**D**) IBD-associated FDPS gene from PsychENCODE eQTLs data; (**E**) CD-associated SULT1A1 protein from discovery pQTLs data; (**F**) CD-associated SULT1A1 gene from PsychENCODE eQTLs data; (**G**) CD-associated FDPS protein from discovery pQTLs data; (**H**) CD-associated FDPS gene from PsychENCODE eQTLs data

among them. Through TWAS secondary validation, only SULT1A1 and FDPS, these two "novel proteins," passed all the aforementioned rigorous criteria, demonstrating strong associations with reduced risk of IBD and CD. These findings represent historical firsts, robustly unveiling new proteins associated with the brain in relation to IBD and CD, potentially serving as critical links between the brain and IBD/CD, or even as potential therapeutic targets.

Currently, the increasing incidence of IBD is partly attributed to rapid urbanization, industrialization, and the Westernization of lifestyle, placing individuals in prolonged high-stress environments. Additionally, dietary changes contribute to the heightened risk of IBD [4]. Furthermore, stress-induced mental disorders may increase

intestinal permeability via the GBA, facilitating the translocation of gut bacteria to peripheral lymphoid organs, triggering innate immune responses, and inducing local gastrointestinal inflammation. Moreover, stress can alter visceral hypersensitivity, motility [28, 29], and involve pathways such as the gut-brain axis, central nervous system responses, and stress pathways (hypothalamic-pituitary-adrenal axis), as well as pathways involving gastrointestinal hormones (e.g., corticotropin-releasing factor) [30], Understanding and managing patients with overlapping disorders of gut-brain interaction), all of which are potential mechanisms promoting the occurrence of IBD. Compared to UC patients, individuals with CD may be more susceptible to psychological factors, exacerbating their condition [31]. This partly explains why this study ultimately identified brain proteins associated with CD rather than UC.

The PFC tissue selected in this study is closely associated with human emotion and mood, which is crucial for chronic stress-induced depressive and anxiety-like behaviors [32]. Prior to this, many key targets and proteins related to emotional disorders have been identified from human brain PFC tissue [33–35]. Based on the concept of the " GBA ", we aimed to explore new targets and proteins associated with IBD to aid in the research and development of new drugs.

Through rigorous testing, including at the protein and mRNA levels, we have identified two proteins, SULT1A1 and FDPS, which consistently influence IBD in the same direction at various levels. SULT1A1 is a member of the human cytosolic sulfotransferase family, catalyzing the metabolism of 3'-phosphoadenosine-5'-phosphosulfate, thereby influencing neurotransmitter transmission [36]. Research indicates that SULT1A1 has two allosteric binding sites, one for catechins (natural polyphenolic compounds) and another for non-steroidal anti-inflammatory drugs, which are believed to be associated with migraine occurrence [37]. This discovery highlights the impact of SULT1A1 on the nervous system. Zhao et al. [38] suggested that SULT1A1 may contribute to improving intestinal inflammation and promoting intestinal barrier repair. Certainly, post-translational modifications (PTMs) of proteins, including neddylation, acetylation, glycosylation, and phosphorylation, also play significant roles in influencing IBD [39]. Ehrentraut SF [40] proposed that PTMs tightly regulate inflammation by altering the functional relevance of protein networks, which is a mechanism relevant to the onset of IBD. The role of SULT1A1 in the interaction between gut microbiota and host metabolism [41] may be influenced by PTMs. Specifically, a critical pathway in gut homeostasis involves the phosphorylation of IκBα and strict regulation of NF-κB activity [42, 43]. Activation of the NF-κB pathway occurs via IκBα phosphorylation, and inhibition of NF-κB activation depends on cullin deneddylation [42]. Notably, neddylation, a PTM, targets cullin proteins specifically [44], facilitating the conjugation of ubiquitin-modified proteins with their target substrates, thereby regulating various cellular processes. Research also indicates that inhibiting cullin 2 neddylation can modulate hypoxia-inducible factor (HIF) to alleviate mucosal inflammation [45, 46].

For FDPS protein, which is involved in the biosynthesis of cholesterol and the metabolism of steroids, and it participates in various crucial cellular functions related to the pathogenesis of tumors, cardiovascular diseases, and autoimmune diseases [47]. However, there is currently no evidence suggesting its association with IBD. Yet, exploring whether FDPS acts as an intermediary in communication between the brain and IBD, along with PTM effects, remains worthwhile. Additionally, this study identified some potential genes, although they did not show satisfactory results in the secondary validation TWAS analysis, they have been previously noted by other scholars. For example, LSP1 has been identified as a risk gene in UC and normal populations [48], consistent with our findings. The deficiency of GPX1 is also considered to be associated with the early inflammatory response of IBD [49]. INPP5E has been identified as a target gene for IBD-associated variations in previous GWAS analyses, and our study further links it to the brain-gut axis.

Despite the rigorous analyses conducted, there are still some noteworthy limitations. Firstly, the limited SNPs in the genetic landscape of brain proteomics stem from the relatively small samples in the original studies, potentially introducing bias into our investigation of the proteome. Similarly, issues arise in transcriptomic studies due to sample size limitations, which could be addressed by developing new databases. Additionally, while this study presents conclusions from both genetic and expressed protein levels, these conclusions are derived from statistical analyses and require further functional studies to validate our findings.

Conclusion

This study provides the first evidence, based on multi-omics data from the brain, supporting the strong associations and causal relationship between two novel proteins, SULT1A1 and FDPS, with IBD and CD, offering new insights into the pathogenesis and potential therapeutic targets for IBD and CD.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12014-024-09511-7.

Supplementary Material 1

Acknowledgements
The authors thank all the participants and researchers who contributed and collected data.

Author contributions
XYF contributed to the conception and design of the study and data analysis. XYF and YZQ contributed to design of the study and wrote the first draft of the manuscript. LLJ contributed to the revision of the manuscript. All authors contributed to the article and approved the submitted version.

Funding
This work is supported by grants from the Jiangxi University of Chinese Medicine Science and Technology Innovation Team Development Program (No. CXTD22011).

Data availability
The ROSMAP and Banner datasets, including weights and pQTL data, can be accessed at https://www.synapse.org/#!Synapse:syn23627957;GWAS summary data for IBD, CD, and UC are sourced from the International IBD Genetics Consortium and can be accessed at https://www.ibdgc.org/;The eQTL summary dataset from PsychENCODE can be found at https://cnsgenomics.com/software/smr/#eQTLsummarydata; The human brain transcriptome data for TWAS is sourced from the CommonMind Consortium and can be accessed at www.synapse.org/CMC.

Declarations

Ethics approval and consent to participate
Not applicable. We used publicly available data that were obtained with ethical approval from their respective institutional review boards and informed consent from all participants. No administrative permissions were required for accessing the data.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 19 June 2024 / Accepted: 30 September 2024
Published online: 16 October 2024

References

1. Ashton JJ, Beattie RM. Personalised therapy for inflammatory bowel disease. Lancet. 2019;393(10182):1672–4.
2. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, Panaccione R, Ghosh S, Wu JCY, Chan FKL, 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.
3. Atreya R, Neurath MF. Biomarkers for personalizing IBD Therapy: the Quest continues. Clin Gastroenterol Hepatol; 2024.
4. Brzozowski B, Mazur-Bialy A, Pajdo R, Kwiecien S, Bilski J, Zwolinska-Wcislo M, Mach T, Brzozowski T. Mechanisms by which stress affects the experimental and clinical inflammatory bowel disease (IBD): role of Brain-Gut Axis. Curr Neuropharmacol. 2016;14(8):892–900.
5. Gracie DJ, Guthrie EA, Hamlin PJ, Ford AC. Bi-directionality of brain-gut interactions in patients with inflammatory bowel disease. Gastroenterology. 2018;154(6):1635–46. e1633.
6. Fairbrass KM, Lovatt J, Barberio B, Yuan Y, Gracie DJ, Ford AC. Bidirectional brain-gut axis effects influence mood and prognosis in IBD: a systematic review and meta-analysis. Gut. 2022;71(9):1773–80.
7. Gao X, Cao Q, Cheng Y, Zhao D, Wang Z, Yang H, Wu Q, You L, Wang Y, Lin Y, et al. Chronic stress promotes colitis by disturbing the gut microbiota and triggering immune system response. Proc Natl Acad Sci U S A. 2018;115(13):E2960–9.
8. Komoto M, Asada A, Ohshima Y, Miyanaga K, Morimoto H, Yasukawa T, Morito K, Takayama K, Uozumi Y, Nagasawa K. Dextran sulfate sodium-induced colitis in C57BL/6J mice increases their susceptibility to chronic unpredictable mild stress that induces depressive-like behavior. Life Sci. 2022;289:120217.
9. Banfi D, Moro E, Bosi A, Bistoletti M, Cerantola S, Crema F, Maggi F, Giron MC, Giaroni C, Baj A. Impact of Microbial metabolites on Microbiota-Gut-Brain Axis in Inflammatory Bowel Disease. Int J Mol Sci 2021, 22(4).
10. Zois CD, Katsanos KH, Kosmidou M, Tsianos EV. Neurologic manifestations in inflammatory bowel diseases: current knowledge and novel insights. J Crohns Colitis. 2010;4(2):115–24.
11. Hall CV, Radford-Smith G, Savage E, Robinson C, Cocchi L, Moran RJ. Brain signatures of chronic gut inflammation. Front Psychiatry. 2023;14:1250268.
12. Ge L, Liu S, Li S, Yang J, Hu G, Xu C, Song W. Psychological stress in inflammatory bowel disease: psychoneuroimmunological insights into bidirectional gut-brain communications. Front Immunol. 2022;13:1016578.
13. Kim JS, Chen MH, Wang HE, Lu CL, Wang YP, Zhang B. Inflammatory bowel disease and neurodegenerative diseases. Gut Liver. 2023;17(4):495–504.
14. Seaton N, Hudson J, Harding S, Norton S, Mondelli V, Jones ASK, Moss-Morris R. Do interventions for mood improve inflammatory biomarkers in inflammatory bowel disease? A systematic review and meta-analysis. EBioMedicine. 2024;100:104910.
15. Graham DB, Xavier RJ. Pathway paradigms revealed from the genetics of inflammatory bowel disease. Nature. 2020;578(7796):527–39.
16. Furey TS, Sethupathy P, Sheikh SZ. Redefining the IBDs using genome-scale molecular phenotyping. Nat Rev Gastroenterol Hepatol. 2019;16(5):296–311.
17. Jia K, Shen J. Transcriptome-wide association studies associated with Crohn's disease: challenges and perspectives. Cell Biosci. 2024;14(1):29.
18. Zhu S, Lin Y, Ding Z. Exploring inflammatory bowel disease therapy targets through druggability genes: a mendelian randomization study. Front Immunol. 2024;15:1352712.
19. Zou M, Liang Q, Zhang W, Zhu Y, Xu Y. Endoplasmic reticulum stress related genome-wide mendelian randomization identifies therapeutic genes for ulcerative colitis and Crohn's disease. Front Genet. 2023;14:1270085.
20. Wingo AP, Liu Y, Gerasimov ES, Gockley J, Logsdon BA, Duong DM, Dammer EB, Robins C, Beach TG, Reiman EM, et al. Integrating human brain proteomes with genome-wide association data implicates new proteins in Alzheimer's disease pathogenesis. Nat Genet. 2021;53(2):143–6.
21. Wingo AP, Fan W, Duong DM, Gerasimov ES, Dammer EB, Liu Y, Harerimana NV, White B, Thambisetty M, Troncoso JC, et al. Shared proteomic effects of cerebral atherosclerosis and Alzheimer's disease on the human brain. Nat Neurosci. 2020;23(6):696–700.
22. Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, Clarke D, Gu M, Emani P, Yang YT et al. Comprehensive functional genomic resource and integrative model for the human brain. *Science* 2018, 362(6420).
23. Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, Ruderfer DM, Oh EC, Topol A, Shah HR, et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat Neurosci. 2016;19(11):1442–53.
24. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, Ripke S, Lee JC, Jostins L, Shah T, 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.
25. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BW, Jansen R, de Geus EJ, Boomsma DI, Wright FA, et al. Integrative approaches for large-scale transcriptome-wide association studies. Nat Genet. 2016;48(3):245–52.
26. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, Timpson NJ, Higgins JPT, Dimou N, Langenberg C, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ. 2021;375:n2233.
27. Storm CS, Kia DA, Almramhi MM, Bandres-Ciga S, Finan C, Hingorani AD, Wood NW, International Parkinson's Disease Genomics C. Finding genetically-supported drug targets for Parkinson's disease using mendelian randomization of the druggable genome. Nat Commun. 2021;12(1):7342.
28. Bonaz BL, Bernstein CN. Brain-gut interactions in inflammatory bowel disease. Gastroenterology. 2013;144(1):36–49.
29. Sun Y, Li L, Xie R, Wang B, Jiang K, Cao H. Stress triggers flare of inflammatory bowel disease in children and adults. Front Pediatr. 2019;7:432.
30. Osadchiy V, Martin CR, Mayer EA. The gut-brain Axis and the Microbiome: mechanisms and clinical implications. Clin Gastroenterol Hepatol. 2019;17(2):322–32.
31. Petruo VA, Krauss E, Kleist A, Hardt J, Hake K, Peirano J, Krause T, Ehehalt R, von Arnauld P, Buning J, et al. Perceived distress, personality characteristics, coping strategies and psychosocial impairments in a national German multicenter cohort of patients with Crohn's disease and ulcerative colitis. Z Gastroenterol. 2019;57(4):473–83.

32. Tong X, Wu J, Sun R, Li H, Hong Y, Liu X, Sun Y, Chen C, Huang L, Lin S. Elevated dorsal medial prefrontal cortex to lateral habenula pathway activity mediates chronic stress-induced depressive and anxiety-like behaviors. Neuropsychopharmacology 2024.
33. Zhang Z, Liu L, Zhang H, Li C, Chen Y, Zhang J, Pan C, Cheng S, Yang X, Meng P, et al. The genetic structure of pain in depression patients: a genome-wide association study and proteome-wide association study. J Psychiatr Res. 2022;156:547–56.
34. Wei W, Zhang H, Cheng B, Qin X, He D, Zhang N, Zhao Y, Cai Q, Shi S, Chu X, et al. Identification of novel functional brain proteins for treatment-resistant schizophrenia: based on a proteome-wide association study. Eur Psychiatry. 2023;66(1):e33.
35. Cheng B, Meng P, Yang X, Cheng S, Liu L, Jia Y, Wen Y, Zhang F. Integrated analysis of proteome-wide and transcriptome-wide association studies identified novel genes and chemicals for vertigo. Brain Commun. 2022;4(6):fcac313.
36. Cook I, Wang T, Leyh TS. Tetrahydrobiopterin regulates monoamine neurotransmitter sulfonation. Proc Natl Acad Sci U S A. 2017;114(27):E5317–24.
37. Vuralli D, Arslan B, Topa E, de Morais AL, Gulbahar O, Ayata C, Bolay H. Migraine susceptibility is modulated by food triggers and analgesic overuse via sulfotransferase inhibition. J Headache Pain. 2022;23(1):36.
38. Zhao XH, Zhao P, Deng Z, Yang T, Qi YX, An LY, Sun DL, He HY. Integrative analysis reveals marker genes for intestinal mucosa barrier repairing in clinical patients. iScience. 2023;26(6):106831.
39. Wang R, Wang G. Protein modification and autophagy activation. Adv Exp Med Biol. 2019;1206:237–59.
40. Ehrentraut SF, Colgan SP. Implications of protein post-translational modifications in IBD. Inflamm Bowel Dis. 2012;18(7):1378–88.
41. Peng B, Zhao H, Keerthisinghe TP, Yu Y, Chen D, Huang Y, Fang M. Gut microbial metabolite p-cresol alters biotransformation of bisphenol A: enzyme competition or gene induction? J Hazard Mater. 2022;426:128093.
42. Neish AS, Gewirtz AT, Zeng H, Young AN, Hobert ME, Karmali V, Rao AS, Madara JL. Prokaryotic regulation of epithelial responses by inhibition of IkappaB-alpha ubiquitination. Science. 2000;289(5484):1560–3.
43. Kelly D, Campbell JI, King TP, Grant G, Jansson EA, Coutts AG, Pettersson S, Conway S. Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. Nat Immunol. 2004;5(1):104–12.
44. Pan ZQ, Kentsis A, Dias DC, Yamoah K, Wu K. Nedd8 on cullin: building an expressway to protein destruction. Oncogene. 2004;23(11):1985–97.
45. Curtis VF, Ehrentraut SF, Campbell EL, Glover LE, Bayless A, Kelly CJ, Kominsky DJ, Colgan SP. Stabilization of HIF through inhibition of Cullin-2 neddylation is protective in mucosal inflammatory responses. FASEB J. 2015;29(1):208–15.
46. MacManus CF, Campbell EL, Keely S, Burgess A, Kominsky DJ, Colgan SP. Anti-inflammatory actions of adrenomedullin through fine tuning of HIF stabilization. FASEB J. 2011;25(6):1856–64.
47. Liu J, Zhang X, Zhang Y, Qian M, Yang M, Yang S, Wang L. Farnesyl diphosphate synthase exacerbates nonalcoholic steatohepatitis via the activation of AHR-CD36 axis. FASEB J. 2023;37(7):e23035.
48. Camarillo GF, Goyon EI, Zuniga RB, Salas LAS, Escarcega AEP, Yamamoto-Furusho JK. Gene Expression Profiling of Mediators Associated with the inflammatory pathways in the intestinal tissue from patients with Ulcerative Colitis. Mediators Inflamm. 2020;2020:9238970.
49. Chu FF, Esworthy RS, Shen B, Gao Q, Doroshow JH. Dexamethasone and Tofacitinib suppress NADPH oxidase expression and alleviate very-early-onset ileocolitis in mice deficient in GSH peroxidase 1 and 2. Life Sci. 2019;239:116884.

Publisher's note