Diagnostic Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Unraveling the genetic susceptibility of irritable bowel syndrome: integrative genome-wide analyses in 845 492 individuals: a diagnostic study

Wentao Huang, MD[a,b], Lijun Zhang, MD[a,c], Yuying Ma, MD[a,b], Shiyi Yu, MD[a,b], Yanlin Lyu, MD[a,d], Shuangshuang Tong, MD[a,d], Jiaxuan Wang, MS[a], Rui Jiang, BS[a,c], Meijun Meng, MD[a,e], Yanjun Wu, MD[a,b], Ruibang Luo, PhD[f,*], Xinqi Qiu, MD[g,*], Weihong Sha, MD[a,b,c,d,e,*], Hao Chen, MD[a,b,c,d,e,*]

Background: Irritable bowel syndrome (IBS) significantly impacts individuals due to its prevalence and negative effect on quality of life. Current genome-wide association studies (GWAS) have only identified a small number of crucial single nucleotide polymorphisms (SNPs), not fully elucidating IBS's pathogenesis.
Objective: To identify genomic loci at which common genetic variation influences IBS susceptibility.
Methods: Combining independent cohorts that in total comprise 65 840 cases of IBS and 788 652 controls, the authors performed a meta-analysis of genome-wide association studies (GWAS) of IBS. The authors also carried out gene mapping and pathway enrichment to gain insights into the underlying genes and pathways through which the associated loci contribute to disease susceptibility. Furthermore, the authors performed transcriptome analysis to deepen their understanding. IBS risk models were developed by combining clinical/lifestyle risk factors with polygenic risk scores (PRS) derived from the GWAS meta-analysis. The authors detect the phenotype association for IBS utilizing PRS-based phenome-wide association (PheWAS) analyses, linkage disequilibrium score regression, and Mendelian randomization.
Results: The GWAS meta-analysis identified 10 IBS risk loci, seven of which were novel (rs12755507, rs34209273, rs34365748, rs67427799, rs2587363, rs13321176, rs1546559). Multiple methods identified nine promising IBS candidate gene (*PRRC2A, COP1, CADM2, LRP1B, SUGT1, MED12L, P2RY14, PHF2, SHISA6*) at 10 GWAS loci. Transcriptome validation also revealed differential expression of these genes. Phenome-wide associations between PRS-IBS and nine traits (neuroticism, diaphragmatic hernia, asthma, diverticulosis, cholelithiasis, depression, insomnia, COPD, and BMI) were identified. The six diseases (asthma, diaphragmatic hernia, diverticulosis, insomnia major depressive disorder and neuroticism) were found to show genetic association with IBS and only major depressive disorder and neuroticism were found to show causality with IBS.
Conclusion: The authors identified seven novel risk loci for IBS and highlighted the substantial influence on genetic risk harbored. The authors' findings offer novel insights into etiology and phenotypic association of IBS and lay the foundation for therapeutic targets and interventional strategies.

Keywords: causality, comorbidity, GWAS meta-analysis, irritable bowel syndrome, polygenic risk scores

[a]*Department of Gastroenterology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University,* [b]*The Second School of Clinical Medicine, Southern Medical University, Guangzhou, China,* [c]*School of Medicine, South China University of Technology,* [d]*Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences,* [e]*Cancer Prevention Center, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou,* [f]*Shantou University Medical College, Shantou, China and* [g]*Department of Computer Science, The University of Hong Kong, Hong Kong*

W.H., L.Z., Y.M., S.Y. contributed equally to this work.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

**Corresponding author. Address: Department of Computer Science, The University of Hong Kong, Hong Kong. Tel.: +852 530 43354. E-mail: rbluo@cs.hku.hk (R. Luo); Cancer Prevention Center, State Key Laboratory of Oncology in South China, Guangdong Provincial Clinical Research Center for Cancer, Sun Yat-sen University Cancer Center, Guangzhou 510060, PR, China. Tel.: +86 177 570 55635. E-mail: qiuxq1@sysucc.org.cn (X. Qiu); Department of Gastroenterology, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, Guangzhou 510080, China. Tel.: +86 138 027 91808. E-mail: shaweihong@gdph.org.cn (W. Sha), and Tel.: +86 150 184 79255. E-mail: chenhao@gdph.org.cn (H. Chen).*

International Journal of Surgery (2025) 111:210–220

Received 18 March 2024; Accepted 30 July 2024

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 14 August 2024

http://dx.doi.org/10.1097/JS9.0000000000002039

Introduction

Irritable bowel syndrome (IBS) typically manifests during early adulthood with symptoms such as abdominal discomfort, bloating, and bowel irregularities[1]. The recently updated Rome IV criteria define IBS as unexplained abdominal discomfort or pain relieved by bowel movements, accompanied by changes in stool consistency or frequency, persisting for more than six months[2]. The prevalence of IBS is estimated to be ~0–15%, making it a leading cause of work-related absenteeism[2]. Almost half of all gastroenterology clinic referrals are due to IBS[3]. Taking into account both direct and indirect expenses associated with IBS, the financial strain is immense, totaling a minimum of €13 billion (£11.7 billion) annually across Europe[4]. Given its high prevalence, significant impact, and economic cost, there is an urgent need for a better understanding of its pathophysiology to facilitate improved therapeutic approaches.

IBS lacking concrete pathology upon clinical or laboratory analysis[2,5]. The diagnosis of IBS is made by identifying characteristic symptoms, as defined by the Rome criteria, and excluding organic gastrointestinal diseases that might otherwise explain these symptoms. An important issue that clinicians face is how to diagnose and treat patients when there exists a co-occurrence between IBS and other diseases. However, the complex associations are unclear. For example, a Swedish population-based case-control study found that the presence of symptomatic uncomplicated diverticular disease was associated with increased abdominal pain and IBSD type symptoms. By contrast, a case-control study from the USA found no association between the presence (or number) of colonic diverticulosis and mucosal inflammation or gastrointestinal symptoms[6]. Hence, a clearer understanding of IBS is essential in identifying the condition when it presents alongside various other symptoms. Addressing these clinically relevant and frequently encountered questions can be facilitated by an enhanced comprehension that IBS is unlikely to be solely a somatosensory disorder[7].

A family study demonstrated that individuals with an affected relative were two to three times to develop IBS[8]. Furthermore, a comprehensive survey across Sweden indicated a heightened risk of IBS in first, second, and third-degree relatives of those with the disease[9]. Heritability estimates derived from twin research show a vast range, spanning from 0 to 57%[8]. As of now, a limited number of genomic regions have been linked to IBS through genome-wide association studies (GWAS)[10–12], leaving much of the heritable component associated with common genetic variations yet to be discovered[13–15]. Identifying genomic risk factors for complex traits like IBS is challenging due to their polygenic architecture, demanding very large sample sizes to identify genetic variants with subtle effects[16]. Thus, larger datasets are essential to fully unravel the heritable components of IBS.

To enhance our understanding of IBS's genetic framework and identify additional genomic locations influencing disease susceptibility, we conducted a meta-analysis of genome-wide association studies (GWAS) of IBS in three independent cohorts that in total comprise 65 840 cases of IBS and 788 652 controls. Subsequently, we integrated gene mapping and pathway enrichments to gain insights into the genetic and biological mechanisms that underlie the significant loci. Finally, we examined the connections between the genetic structure of IBS and various other traits.

HIGHLIGHTS

- We using the largest dataset to perform GWAS meta-analysis. A total of 10 risk single nucleotide polymorphisms (SNPs) were identified for IBS and seven risk SNPs were firstly identified (rs12755507, rs34209273, rs34365748, rs67427799, rs2587363, rs13321176, rs1546559).
- Nine functional genes associated with IBS (*PHF2*, *MED12L*, *LRP1B*, *COP1*, *PRRC2A, MED12L*, *P2RY14*, *CAMD2*) were identified. The knock-out model provided evidence for the associations between IBS and phenotypic changes in the respiratory, digestive, nervous systems, and lipid metabolism.
- Our study performed PRS-based PheWAS analysis and identified phenotypic associations between IBS and nine traits (neuroticism, diaphragmatic hernia, asthma, diverticulosis, cholelithiasis, depression, insomnia, COPD, and BMI).
- Mendelian Randomization analysis provided a more accurate demonstration of the causality between IBS and major depressive disorder, as well as neuroticism, and corrected the biases observed in previous studies concerning the association of IBS with other diseases, including diaphragmatic hernia, diverticulosis, asthma, and COPD, using the most strongly associated SNPs.

Methods

Study design

The design flow chart is shown in Figure 1. The work has been reported in accordance with the STROCSS, Supplemental Digital Content 1, http://links.lww.com/JS9/D336 criteria[17] and in line with the STARD, Supplemental Digital Content 2, http://links.lww.com/JS9/D337 (Standards for the Reporting of Diagnostic accuracy studies) criteria.

Cohorts

Data were acquired from three distinct datasets, totaling 65 840 cases of IBS and 788 652 control subjects in the primary GWAS. We combined GWAS summary data from Finngen, UKB/Bellygenes, and the Resource for Genetic Epidemiology Research on Aging (GERA). The diagnosis of IBS in these cohorts is based on ICD-9 or ICD-10 codes. Specifically, the Finngen project identified IBS cases using hospital or registry records with ICD-9 or ICD-10 coding. In the UKB/Bellygenes cohort, IBS cases were identified based on digestive health questionnaire (DHQ) Rome III symptom data, self-reported previous medical IBS diagnoses, or electronic medical records. Data for the GERA cohort were derived from ICD-9 coded diagnoses in Kaiser Permanente's electronic medical records (Supplementary Table S1, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). Post-regression quality control (QC) procedures were conducted using EasyQC[18]. In this process, datasets were compared to the 1000 Genomes reference panel, excluding variants showing an allele frequency deviation beyond 0.2 when compared to the reference. A meta-analysis was performed, adopting the fixed-effect inverse-variance-weighted approach with METAL[19], including only variants appearing in a minimum of two datasets out of three in

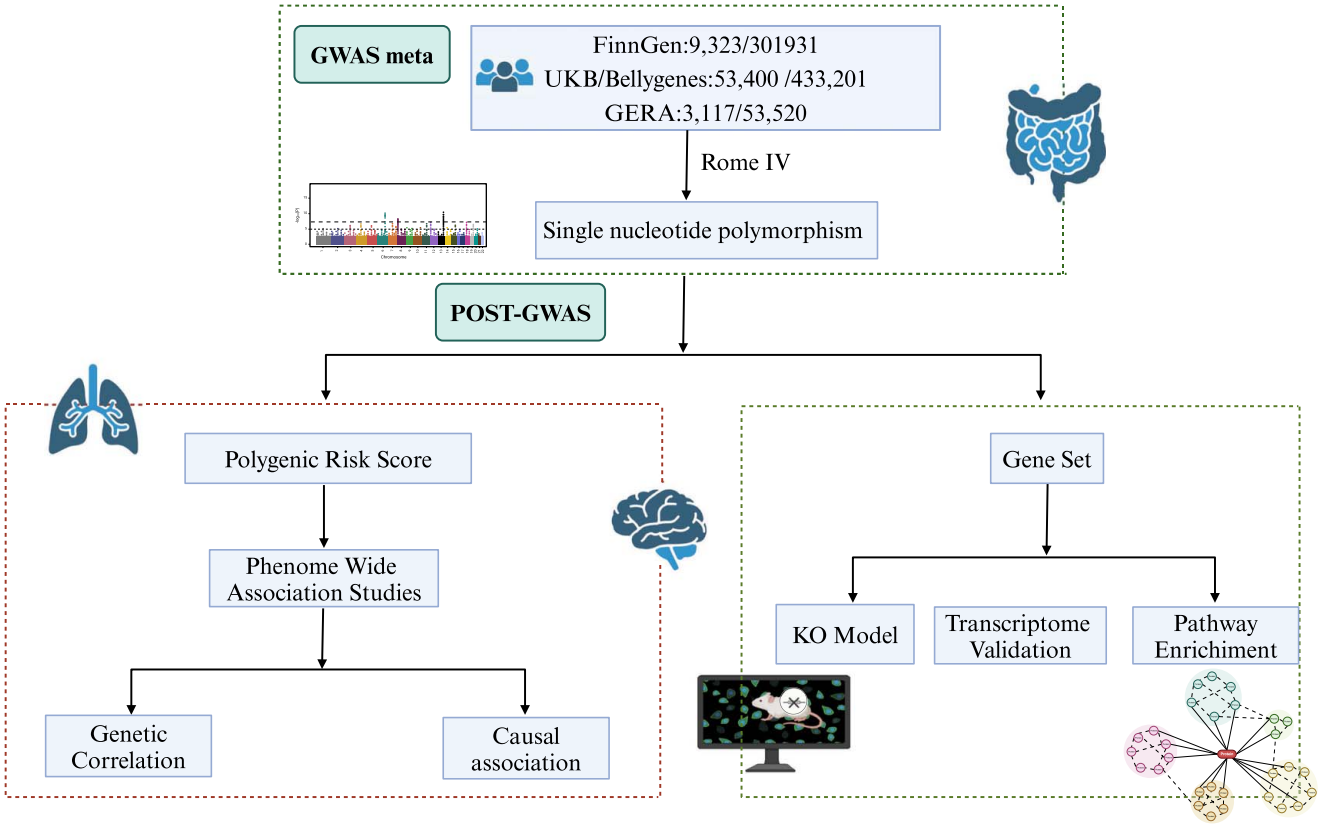

Figure 1. Study design and workflow. GWAS, genome-wide association studies.

the final meta-analysis. We computed genomic inflation factors and LDSC intercept calculations for every individual cohort and the comprehensive meta-analysis[20]. We plotted the Manhattan to visualize the results using R package "CMplot"[21].

Locus definition/LD clumping

Linkage disequilibrium (LD) clumping was used to identify the positions of loci containing IBS-associated variants. We identify independently significant SNPs using this thread: 1 Mb window (–clump-kb 1000) and a stringent LD threshold (–r^2 0.001). We set the significance threshold at $P < 5 \times 10^{-8}$).

Gene mapping

We employed three primary methods for gene mapping. Firstly, PoPS, a gene prioritization tool reliant on similarity, was engaged to discern potential causal genes, integrating GWAS summary statistics, gene expression data, biological pathways, and predicted protein-protein interaction insights from a comprehensive set of over 50 000 features[22]. For gene prioritization with PoPS, gene-level association statistics and correlations amongst genes were initially calculated, drawing from our IBS summary statistics and employing MAGMA[23], followed by gathering LD estimates from 1000 Genomes European samples. We subsequently performed enrichment analyses of gene features. Finally, PoPS scores were calculated for each gene, fitting a joint model for all feature enrichments, and potential candidate genes need PoPS z-scores surpassing 0.25. When two genes within the same genomic locus both had z-scores above 0.25, priority was given to the gene with the highest PoPS score. Secondly, we traced lead variants or those in significant linkage disequilibrium ($r^2 > 0.8$) anticipated to impact protein-coding or splicing, using Variant Effect Predictor (VEP)[24]. Thirdly, we linked lead variants from each risk locus to genes utilizing eQTL data derived from GTEx version 8, covering tissues like nerve, lung, and whole blood, using Functional Mapping and Annotation of Genetic Associations (FUMA) application[25,26].

Pathway analysis

The MAGMA analyses were performed using the default settings of FUMA[23,25] to conduct pathway analyses, employing data from 1000 Genomes Phase 3 as the LD reference and incorporating 15,485 pathways from MsigDB version 7.0[27].

Querying the MGI database

We accessed the Mouse Genome Informatics (MGI) resource at http://www.informatics.jax.org/ to retrieve all candidate genes from our novel GWAS hits list. MGI employs standardized nomenclature and controlled vocabularies, which include the Mouse Developmental Anatomy Ontology, the Mammalian Phenotype Ontology, and the Gene Ontologies. Since MGI compiles and organizes data from primary literature sources, we systematically reviewed all reported systemic abnormalities associated with models related to the queried genes[28]. To gain a deeper understanding of the association between IBS and other phenotypes, we conducted manual curation of these abnormalities.

Transcriptome validation

We selected bulk sequencing data from the GEO dataset (GSE36701), which includes data from 18 constipation-predominant IBS subjects (IBS-C), 27 diarrhea-predominant IBS subjects (IBS-D), 21 individuals with Campylobacter jejuni infection, and 40 healthy volunteers (HV)[29]. Then, we visualized the results of the differential analysis using R with the "ggplot2" (https://ggplot2.tidyverse.org) and "ggpubr" (https://rpkgs.datanovia.com/ggpubr/) packages.

PRS calculations

For the construction of an Polygenic Risk Score (PRS) for IBS, a meta-analysis was performed excluding the UKB as our discovery sample. This analysis resulted in 29 270 cases and 383 647 controls. Our main aim was to pinpoint the effect magnitude of genetic variants, which showed a nominal association to the disease, based on an established *P* value and allele risk weighting. We derived our target sample from the UKB, consisting of 412 917 individuals. We excluded those who were outliers due to heterozygosity, had a variant call rate less than 99%, or demonstrated relatedness up to the third degree with kinship coefficients greater than 0.044, as computed with the ukbtools packages in R[30]. Additionally, we ruled out variants exhibiting a call rate of less than 99%, a minor allele frequency under 0.01, deviation from Hardy-Weinberg equilibrium with *P* value less than 10^{-6}, or those on an ambiguous strand. After employing LD clumping with a threshold of r^2 less than 0.1 in a 250 kb window, using the previously mentioned LD reference, 12 462 SNPs remained that demonstrated relative independence for PRS calculation.

In our target sample, we computed PRS for IBS as the sum of imputed SNP dosages weighted by allele effects (logistic regression coefficients) from the discovery set across all SNPs, employing *P* value thresholds of 1×10^{-3}, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5, respectively. PRS profiling was performed using PLINK (version 1.9)[31,32]. All scores were standardized within each target set to account for variations in the number of SNPs used for PRS calculation. Subsequently, we assessed associations using logistic regression in R (version 4.2.0). Starting from a low *P* value threshold moving up to *P* value 0.5, an optimal *P* value threshold with the highest explained variance was identified calculated by comparing the variance of a full model that included the PRS-IBS with a baseline model that consisting of sex, age, and the first four PCs[33].

PRS evaluation

Firstly, validate if these findings are applicable to IBS, a comparative analysis was conducted between individuals with a PRS exceeding the 90th percentile and the rest within the score distribution. This analysis encompassed 29 270 IBS patients and 383 647 control participants from the UKB, employing logistic regression while adjusting for variables including age, sex, and the initial four Principal Components (PCs). Secondly, we estimated the incremental variance explained (r^2) by the PRS-IBS and compared it to a PRS based on the five lead SNPs. The difference between comprehensive model (including PRS, gender, age, and four PCs and a null model (containing only age, gender, and four PCs) were detected, after we estimated the variance explained on the observed scale by calculating Nagelkerke's r^2. Additionally, we assessed the potential improvement in participant discrimination and reclassification by the PRS-IBS. The discriminative ability of PRS-IBS was compared with non-genetic risk factors associated with IBS, using the AUC metric. After iteratively incorporating risk factors, we quantified the change in AUC relative to a baseline model that included age, sex, and the first four PCs. The calculation of the area under the curve (AUC) was performed using the pROC package in R[33].

PRS-based PheWAS

We conducted a phenome-wide association analyses (PheWAS) for the PRS-IBS with 27 traits, concentrating on diseases pertaining to the digestive tract, mental health, cardiovascular system, and respiratory organs. Logistic regression was performed to analyze. All models were adjusted for sex, age and the first four PCs. To avoid false positives, FDR correction was performed[34] and a corrected FDR-*P* value less than 0.05 was considered to be statistically significant.

Genetic correlation

Genetic correlation estimates using LDSC (V.1.0.1)[20,35] to analyze the genetic correlation between IBS and those positive binary and quantitative traits in the PheWAS. Given that many of the examined traits incorporated data from the UK Biobank, we conducted a meta-analysis for all our IBS cohorts, excluding the UKB dataset, to ensure non-overlapping samples. For the analysis of other traits, we utilized data from the UK Biobank. The provided LD scores were used, as they are appropriate for the European population. All parameters were used in default settings.

Mendelian randomization

Mendelian randomization (MR) was used to infer causality for those positive results in PheWAS following the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement[36].To prevent overlap and ensure larger sample sets, we adopted several strategies. First, we deliberately selected different datasets for each outcome to avoid duplication. Second, for traits exclusively available in the UK Biobank cohort, we conducted a meta-analysis of all IBS cohorts, excluding the UK Biobank sample set. Third, for traits only present in the GERA cohort, we performed a meta-analysis of all our IBS cohorts, except the GERA sample set. Genetic instruments were identified from the papers referenced above and used in LDSR. The R package "TwoSampleMR"[37] was used to perform the analyses. We set the significance threshold ($P \leq 5 \times 10^{-8}$) to get the SNPs for IBS. If the SNPs less than 2 SNPs to perform inverse variance-weighted model, we relaxed the significance threshold ($P \leq 5 \times 10^{-6}$) to make sure get the enough SNPs for IBS. Subsequently, we calculated F-statistics ($F = \text{beta.exposure}^2/\text{se.exposure}^2$), selecting only those with F greater than 10. In our foundational analyses leveraging the inverse variance-weighted model, we established a significance threshold for associations at *P* less than 5.56×10^{-3}, whereas in our weighted median sensitivity analyses, a *P* value under 0.05 was considered significant. The pleiotropy were detect MR–Egger–intercept ($P \geq 0.05$) and "MR-PRESSO"[38]. We also removed genes that suggested greater variance than exposure using Steiger filtering.

Results

GWAS risk loci

We selected SNPs that were present in at least two GWAS studies, ultimately resulting in a new GWAS summary including 7 322 892 SNPs. The LDSC intercept was 1.0179 (se = 0.0085), suggesting that most of the inflation is due to IBS polygenicity (Supplementary Table S2, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). We observe association at three previously reported IBS susceptibility loci (rs2736155, rs1248825, rs10156602) and seven novel IBS susceptibility loci (rs12755507, rs34209273, rs34365748, rs67427799, rs2587363, rs13321176, rs1546559) using these thread P less than 5×10^{-8} (Fig. 2, Table 1, Supplementary Table S3, Supplemental Digital Content 3, http://links.lww.com/JS9/D338).

Gene set

We used three main methods to identify the potential genes: PoPS, Variant Effect Predictor and eQTL. Using PoPS (Supplementary Table S4, Supplemental Digital Content 3, http://links.lww.com/JS9/D338), we initially conducted Pops analysis and identified five genes (*PHF2*, *MED12L*, *LRP1B*, *COP1*, *PRRC2A*) using the human tissue panel dataset with PoPS z >0.25. In addition, the VEP analysis identified nine genes associated with these SNPs. Notably, the SNPs rs2736155 and rs13321176 were found to be associated with multiple genes (Supplementary Table S5, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). We also mapped 12 genes with eQTL (including gut-related tissues and other reported associated tissues) to get more stability association with genes (Supplementary Table S6, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). Ultimately, synthesizing evidence from these three lines of analysis, we identified nine candidate genes including *PRRC2A, COP1, CADM2, LRP1B, SUGT1, MED12L, P2RY14, PHF2, SHISA6* (Table 1).

We performed the gene set enrichment analysis to gain insight into the biological mechanisms, particularly their influence on neurosystem-related biological processes such as the negative regulation of embryonic development, regulation of multicellular organismal development, migration of glial cells in the telencephalon, development of the cerebral cortex. (Supplementary Table S7, Supplemental Digital Content 3, http://links.lww.com/JS9/D338).

Mouse knock-out models for novel genes identified by GWAS

We queried for knock-out (KO) mouse models, using the Mouse Genomics (MGI) resource, for evidence that modification of the target produces a phenotype relevant to IBS. In nine potential risk genes, we retrieved evidence for the association with the gut-organ axis such as decreased airway responsiveness, abnormal brain morphology, decreased white fat cell number, impaired spatial learning and impaired gastric peristalsis, suggesting an intrinsic role in the gut-organ Axis (Supplementary Table S8, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). Additionally, transcriptome validation provided further evidence supporting the functional relevance of our identified genes (Supplementary Figure S1, Supplemental Digital Content 3, http://links.lww.com/JS9/D338).

PRS and IBS risk prediction

The PRS-IBS comprised approximately 12 462 SNPs and underwent association testing in a cohort of 29 270 IBS cases and 383 647 controls from the UK Biobank. To contextualize the PRS-IBS in a clinical context, we compared the risk associated with different levels of high PRS (>90th percentile) against the remainder. Individuals in the highest PRS-IBS decile exhibited an elevated risk of developing IBS, with an odds ratio (OR) of 1.105 (95% CI 1.063–1.149, $P = 3.20 \times 10^{-7}$). Additionally, we observed that the 90th percentile of the PRS, using the lead SNPs associated with IBS risk, was also linked to an increased risk of

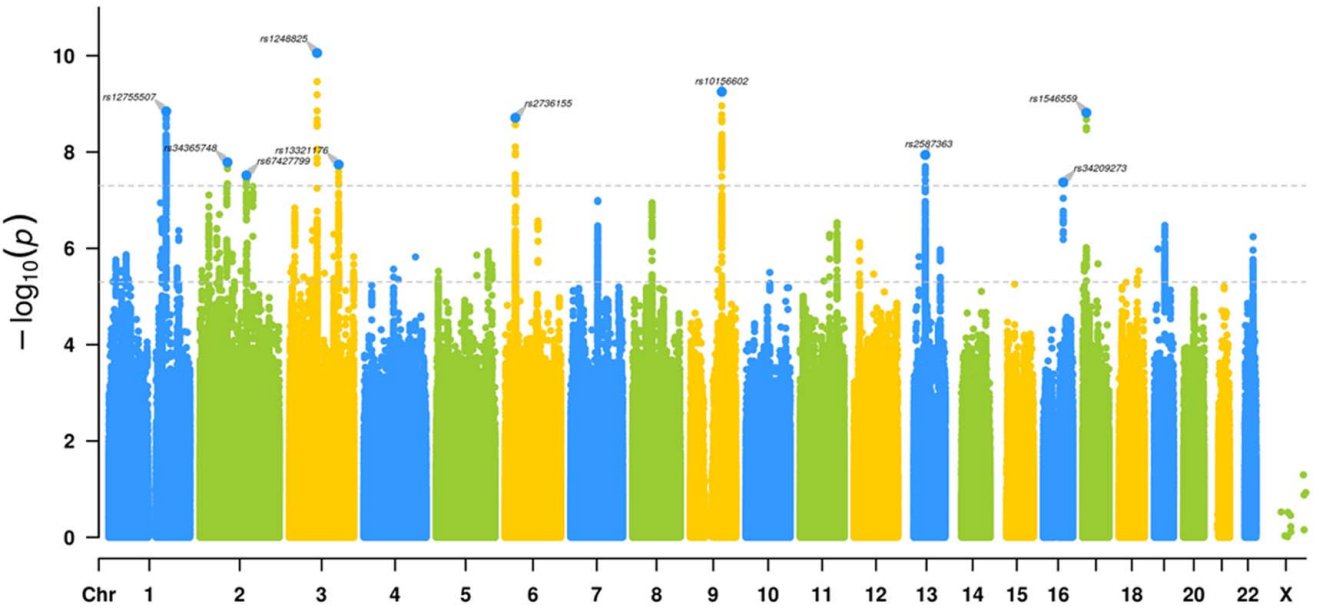

Figure 2. Manhattan plot showing genome-wide significant loci associated with IBS (65 840 cases of IBS and 788 652 control). Genome-wide significance threshold ($P = 5 \times 10^{-8}$).

Table 1

Ten genetic loci associated with irritable bowel syndrome.

SNP	Chromosome	Position[a]	Gene[b]	EAF	Meta-analysis						
					Beta	SE	*P*	Direction	Allele1	Allele2	Reported
rs2736155	6	31605199	PRRC2A	0.504992	0.0382	0.0064	1.95E-09	+?+	C	G	PMID: 34741163
rs12755507	1	176164865	COP1 (also called RFWD2))	0.724241	0.0385	0.0064	1.43E-09	+++	T	C	
rs34209273	16	60721642	NA	0.282947	−0.0339	0.0062	4.25E-08	-?-	A	G	
rs1248825	3	84993411	CADM2	0.307708	0.0426	0.0066	8.79E-11	+?+	A	C	PMID: 34741163
rs34365748	2	84275796	NA	0.0940495	0.0404	0.0072	1.63E-08	+++	T	C	
rs67427799	2	142181419	LRP1B	0.154553	−0.0437	0.0079	3.05E-08	—	A	C	
rs2587363	13	53915933	SUGT1	0.6877	−0.0358	0.0063	1.15E-08	-?-	A	G	
rs13321176	3	150944808	MED12L/P2RY14/	0.9069489	0.0466	0.0083	1.82E-08	+++	T	G	
rs10156602	9	96345328	PHF2	0.696286	0.0399	0.0064	5.62E-10	+?+	A	G	PMID: 34741163
rs1546559	17	11228954	SHISA6	0.489018	−0.0374	0.0062	1.53E-09	-?-	C	G	

Lead variants at each independent locus, including heterogeneity test and direction of effect. Plus sign (+) indicates a positive direction of effect and a minus sign (-) indicates a negative direction of effect. Question mark means that the variant is missing in the specific cohort.
Effect estimates (Beta), standard errors and two-sided *P* values were obtained from inverse-variance-weighted fixed-effects meta-analysis, and not adjusted for multiple-testing
Allele1, effect allele; Allele2, non-effect allele; EAF, effect allele frequency; *P*, *P* value; SE, standard error.
[a]GRCh37 assembly.
[b]Nearest protein-coding gene.

IBS (OR = 1.076, 95% CI 1.035–1.118, $P = 0.0002$). These findings remained consistent when adjusting for ten or twenty PCs compared to a model with four PCs (Supplementary Table 9, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). We estimated phenotypic variance explained (r^2) for the new PRS-IBS relative to a 5-SNP PRS derived from lead SNPs from this study. We observed that the PRS-IBS explained the higher proportion of phenotypic variation (Supplementary Table 10, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). By analyzing the supplementary predictive capacity of distinct demographic and clinical risk factors for IBS, we determined that the inclusion of PRS-IBS enhanced IBS prediction in a model that considered all other IBS risk factors (AUC-risk factors = 0.6987 vs. AUC-risk factors + PRS = 0.6990, delta-AUC + 0.0012, 95% CI 0.00008–0.0004, $P = 0.006$).

Phenotypic and causal relationship with selected traits

PheWAS results for the PRS-IBS and 27 traits were shown in Figure 3A **and** Supplementary Table S11, Supplemental Digital Content 3, http://links.lww.com/JS9/D338. Notable associations were observed in neuroticism, diaphragmatic hernia, asthma, diverticulosis, cholelithiasis, depression, insomnia, COPD, and BMI. We further detected the genetic correlation using LDSC and found strong genetic correlation between IBS and neuroticism, diaphragmatic hernia, asthma, diverticulosis, depression and insomnia (Fig. 3B**and** Supplementary Table S12, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). Finally, we revealed the causality between IBS and neuroticism and depression using mendelian randomization (Fig. 4 and Supplementary Table S13, Supplemental Digital Content 3, http://links.lww.com/JS9/D338). All GWAS summary data details are in Supplementary Table S14, Supplemental Digital Content 3, http://links.lww.com/JS9/D338.

Discussion

The present study is the largest GWAS meta-analysis of IBS concluding 65 840 cases of IBS and 788 652 controls, identifying 10 IBS risk loci, seven of which are entirely novel. Combing PRS-based PheWAS analysis, we detected the underlying pleiotropic associations between IBS and nine traits, which might provide new biological insights in shared genetic architecture. Further, we found more stable causality between genetically predicted IBS with these traits. Therefore, our findings suggest a focus on the comorbidity of IBS, including the diagnosis and treatment of these comorbidities. In discussion,

In the seven newly identified IBS risk loci, we found seven candidate genes significantly associated with IBS. An altered crosstalk between gut microbiota and the immune system has been discussed in IBS[39]. *SUGT1*(suppressor of G2 allele of skp1, rs2587363) was identified with the intestinal antibacterial gene and more highly expressed in immune normal IBS patients compared to healthy subjects[40], which is important for chronic low-grade inflammation, and encoding AMPs are lacking in the context of IBS[41]. On the other hand, the constipated symptom is a typical subtype of irritable bowel syndrome, and its association with colorectal cancer (CRC) remains a controversial issue[42–45]. In our research, many potential risk genes for IBS, which were mapped to SNPs, were reported to be associated with CRC. For instance, rs12755507 was mapped to the ubiquitin ligase *COP1*

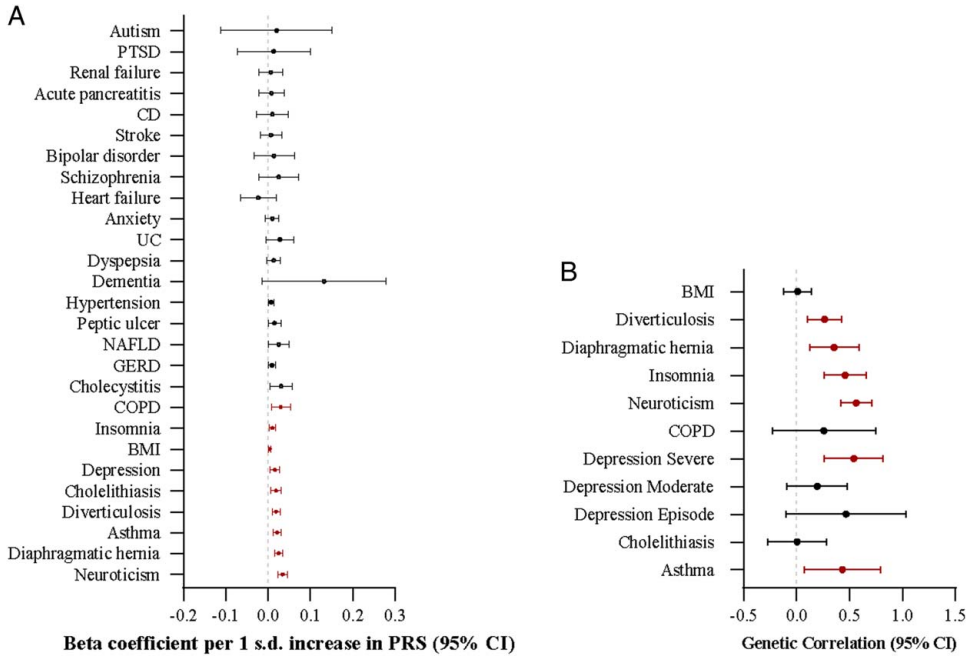

Figure 3. Phenome-wide associations and genetic correlation. (A) Phenome-wide associations between PRS-IBS and selected traits. (B) Genetic correlation between IBS and selected traits. CD, Crohn's disease; COPD, chronic obstructive pulmonary disease; GERD, gastroesophageal reflux disease; IBS, irritable bowel syndrome; NAFLD, non-alcoholic fatty liver disease; PTSD, post-traumatic stress disorder; UC, ulcerative colitis.

(Constitutive Photomorphogenic 1), which acts as a tumor suppressor through the negative regulation of ETV1 (ETS Variant Transcription Factor 1)[46]. Another study reported substantial differences in the expression level of *MED12* (Mediator Complex Subunit 12, rs13321176) among colorectal cancer cell lines and demonstrated an inverse correlation with transforming growth factor (TGF)-β signaling and a direct correlation with apoptosis in response to chemotherapeutic agents[47]. Additionally, *LRP1B* (Low Density Lipoprotein Receptor-Related Protein 1B, rs67427799) and *P2RY14* (Purinergic Receptor P2Y, G-Protein Coupled, 14, rs13321176) were identified as potential prognostic biomarkers for CRC[48,49]. Therefore, these results suggest the need for further research into the relationship between constipation-predominant IBS and CRC. The relationship between IBS and the nervous system was also a critical aspect of the pathogenesis and progression of IBS. We detected *SHISA6* (Shisa Family Member 6, rs1546559) as a potential risk gene for IBS, associated with neural cell growth and development and glioma[39–41]. We successfully replicated earlier findings of three risk loci (rs10156602, rs34209273, rs2736155) for IBS, initially identified in a genome-wide analysis of 53,400 individuals[12]. Consistent with previous findings, these SNPs were associated with three risk genes *CADM2* (Cell Adhesion Molecule 2), *PHF2* (Plant Homeodomain Finger 2) and *BAG6* (BCL2-Associated Athanogene 6), linked to anxiety disorders[50,51], neuroticism, depression, and autism spectrum disorder[52,53]. However, diverging from previous findings, in our analysis, we mapped the gene *PRRC2A* (proline-rich coiled-coil 2A) to rs2736155. *PRRC2A* has been reported to be enriched in the colon tissues[54,55]. and is associated with various autoimmune diseases, including multiple sclerosis (MS)[56]. All these results revealed the important role of Gut-Brain axis in the development of IBS and psychiatric disorders, particularly depression and neuroticism[57,58].

Our research identified strong associations with multiple traits based our novel GWAS, establishing a more stable phenotype association. In our KO model, in addition to displaying various neurological anomalies such as altered brain development, structure, and reduced oligodendrocyte progenitor cells, we discovered additional phenotype associations for these genes. These included decreased airway responsiveness, reduced mesenteric fat pad weight, decreased subcutaneous adipose tissue amount, abnormal adipose tissue development, and other obesity-related phenotypes.

Furthermore, we conducted a PRS-based PheWAS analysis to identify associations between genetic liability for IBS and 27 traits available in the UK Biobank dataset. We were able to confirm previously observed associations of various phenotypes with IBS, including neuroticism[59], anxiety[60] and depression[59]. Additionally, the association we observed between IBS and respray diseases including COPD[61] and asthma[62], was also previously observed in a prospective cohort study. Moreover, a Mendelian randomization study found a positive association between genetically predicted childhood-onset asthma and IBS[63]. However, the association with IBS was based on only two significant SNPs, which may yield uncertain results. After performing a large-scale GWAS meta-analysis, we identified additional SNPs and found evidence suggesting that the association between asthma, COPD and IBS might be due to shared genetic architecture.

We also observed a difference in the association of IBS and both diaphragmatic hernia and diverticulosis. We found a strong genetic correlation, and the results of our PRS-based PheWAS analysis indicated an association between IBS and both conditions. However, no causality was identified. Li *et al.*[57] found

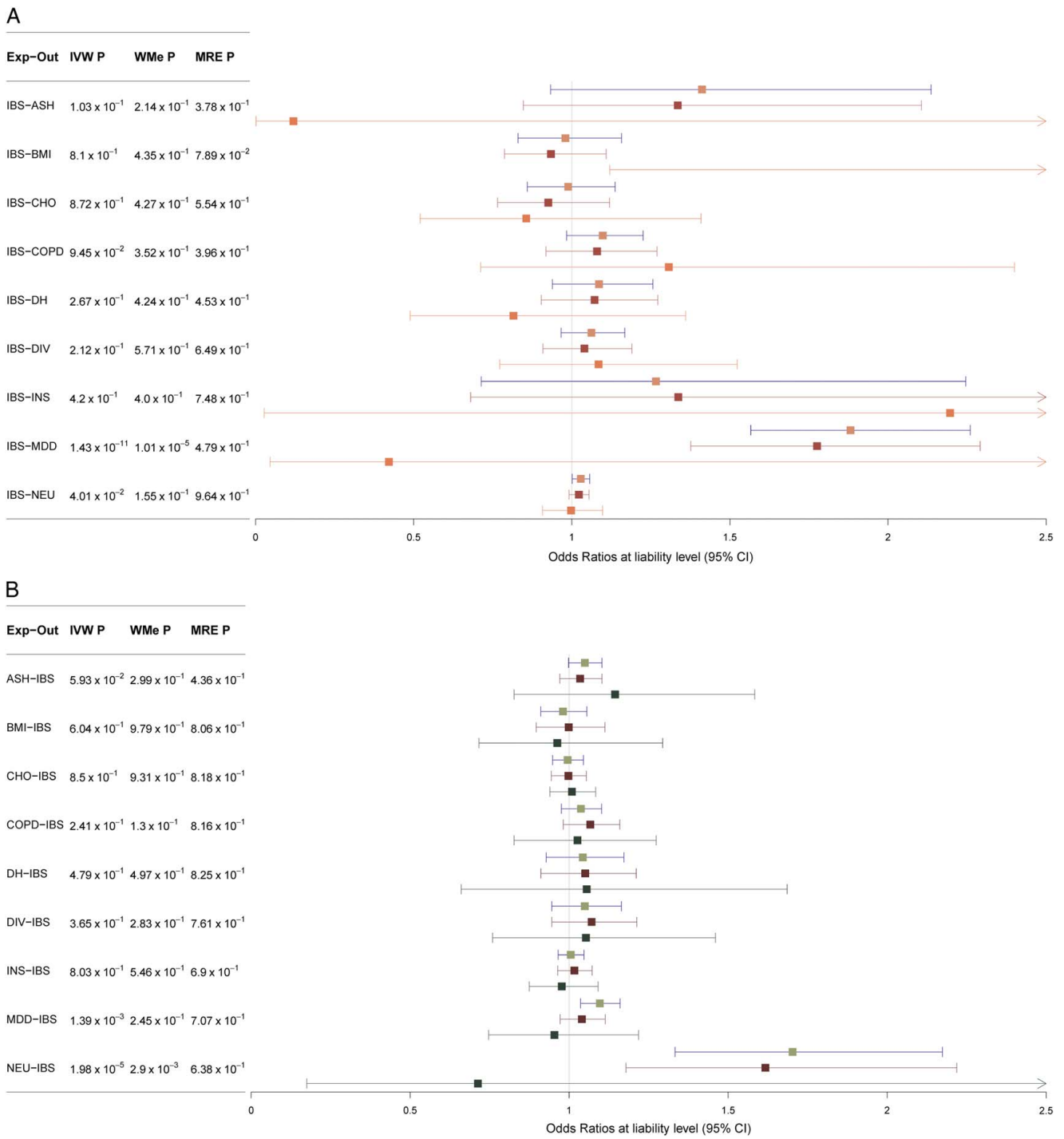

Figure 4. Causal association between IBS and selected traits. (A) The causal effect of genetically predicted IBS on nine traits. For each trait, three boxes represent IVW. (B) The causal effect of nine traits on genetically predicted IBS. For each trait, three boxes represent. ASH, asthma; CHO, cholelithiasis; COPD, chronic obstructive pulmonary disease; DH, diaphragmatic hernia; DIV, diverticulosis; IBS, irritable bowel syndrome; INS, insomnia; IVW, inverse variance weighting, MR Egger, Weighted Median; MDD, major depressive disorder; NEU, neuroticism.

bidirectional causal associations between IBS and diaphragmatic hernia. However, significant heterogeneity among the instrumental variables was observed, and the number of IBS cases was limited. Therefore, our results suggest that patients with diaphragmatic hernia and diverticulosis might not benefit from preventive treatment for IBS.

Additionally, we observed a significant association between BMI and IBS. But we failed to find more evidence for the potential genetic association. More functional research would need to be performed to find the association between obesity and IBS.

The limitations of the current analysis are multiple. Firstly, the sample population consists only of individuals of European

descent, which may limit the generalizability of our findings to other racial groups. Secondly, we were unable to consider the influence of gender factors due to the lack of gender stratification in the GWAS data. Thirdly, the limited availability of GWAS data for neuroticism, insomnia, diaphragmatic hernia, and diverticulosis prevented us from including complete GWAS-meta results for each phenotype, necessitating the exclusion of IBS from the same cohort and the need for a new meta-analysis. Therefore, to obtain further confirmation of the relationship between individual phenotypes and IBS, larger-scale GWAS data is required.

In summary, 10 genetic risk loci were identified in this GWAS meta-analysis of the largest IBS cohort available to date. PRS-based PheWAS IBS led to the identification of associations with nine phenotypes and revealed the potential genetic pleiotropy. The causal relationships were detected between IBS and MDD and neuroticism. These findings shed light on the genetic mechanisms underlying IBS and its potential comorbidity.

Ethical approval

All data used in our research were obtained with informed consent from participants and ethical approval. No additional ethical approval was required as no new data were collected.

Consent

Not applicable. Because, no additional ethical approval was required as no new data were collected.

Source of funding

This work was funded by National Natural Science Foundation of China Regional Innovation and Development Joint Foundation (U23A20408), the National Natural Science Foundation of China (82171698, 82170561, 81300279, and 81741067), the Natural Science Foundation for Distinguished Young Scholars of Guangdong Province (2021B1515020003), Natural Science Foundation of Guangdong Province (2022A1515012081), the Climbing Program of Introduced Talents and High-level Hospital Construction Project of Guangdong Provincial People's Hospital (DFJH201803, KJ012019099, KJ012021143, and KY012021183).

Author contribution

Conceptualization and design: X.Q.Q., R.B.L., W.H.S., and H.C. Funding acquisition: W.H.S. and H.C. Collecting and assembly the data: W.T.H., Y.Y.M., L.J.Z. Data analysis and interpretation: W.T.H., L.J.Z., Y.Y.M. Visualization and validation: L.J.Z., J.X.W., Y.Y.M., S.S.T., Y.L.L., M.J.M., Y.J.W. and R.J. Manuscript writing: W.T.H., Y.Y.M., L.J.Z. and H.C. All authors reviewed and approved the final manuscript.

Conflicts of interest disclosure

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Research registration unique identifying number (UIN)

Not applicable. Because, no additional ethical approval was required as no new data were collected. All data in our research was public bioinformatic data.

Guarantor

Ruibang Luo, Xingqi Qiu, Weihong Sha, Hao Chen.

Data availability statement

Data source: GWAS summary statistics for IBS are available by application from:
Finngen: https://www.finngen.fi/en/access_results.
GERA: http://cg.bsc.es/gera_summary_stats/.
UKB & Bellygenes: http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90016001-GCST90017000/GCST90016564/.
Code availability: The following software and packages were used for data analysis.
PLINK v.2.0 (https://www.cog-genomics.org/plink/2.0/),
METAL v.2011-0325 (http://csg.sph.umich.edu/abecasis/Metal/download/),
MAGMA v.1.07 (https://ctg.cncr.nl/software/magma),
EasyQC v.9.2 (https://www.uni-regensburg.de/medizin/epidemiologie-praeventivmedizin/genetische-epidemiologie/software/),
FUMA v.1.3.8 (https://fuma.ctglab.nl/),
LD score regression v.1.0.1 (https://github.com/bulik/ldsc),
PoPS v.0.1 (https://github.com/FinucaneLab/pops/tree/add-license-1),
TwoSampleMR v.0.5.6 (https://mrcieu.github.io/TwoSampleMR/), and R v.4.1.2 (https://www.r-project.org/).

Provenance and peer review

Not commissioned, externally peer-reviewed.

Acknowledgement

The authors gratefully acknowledge the contributions from public available databases and the participants who contributed to those studies.

References

[1] Ac F. Commentary: estimating the prevalence of IBS globally-past, present and future. Aliment Pharmacol Therap 2020;51:198–9.
[2] Gf L, Wg T, Wd C, *et al*. Functional bowel disorders. Gastroenterology 2006;130:1480–91.
[3] Soubieres A, Wilson P, Poullis A, *et al*. Burden of irritable bowel syndrome in an increasingly cost-aware National Health Service. Frontline Gastroenterol 2015;6:246–51.
[4] Hillilä MT, Färkkilä NJ, Färkkilä MA. Societal costs for irritable bowel syndrome--a population based study. Scand J Gastroenterol 2010;45:582–91.
[5] Camilleri M. Diagnosis and treatment of irritable bowel syndrome: a review. JAMA 2021;325:865–77.
[6] Peery AF, Keku TO, Addamo C, *et al*. Colonic diverticula are not associated with mucosal inflammation or chronic gastrointestinal symptoms. Clin Gastroenterol Hepatol 2018;16:884–891.e1.

[7] Ford AC, Lacy BE, Talley NJ. Irritable bowel syndrome. N Engl J Med 2017;376:2566–78.
[8] Saito YA. The role of genetics in IBS. Gastroenterol Clin North Am 2011; 40:45–67.
[9] Waehrens R, Ohlsson H, Sundquist J, *et al*. Risk of irritable bowel syndrome in first-degree, second-degree and third-degree relatives of affected individuals: a nationwide family study in Sweden. Gut 2015;64:215–21.
[10] Holliday EG, Attia J, Hancock S, *et al*. Genome-wide association study identifies two novel genomic regions in irritable bowel syndrome. Am J Gastroenterol 2014;109:770–2.
[11] Bonfiglio F, Zheng T, Garcia-Etxebarria K, *et al*. Female-specific association between variants on chromosome 9 and self-reported diagnosis of irritable bowel syndrome. Gastroenterology 2018;155:168–79.
[12] Eijsbouts C, Zheng T, Kennedy NA, *et al*. Genome-wide analysis of 53,400 people with irritable bowel syndrome highlights shared genetic pathways with mood and anxiety disorders. Nat Genet 2021;53: 1543–52.
[13] Bonfiglio F, Henström M, Nag A, *et al*. A GWAS meta-analysis from 5 population-based cohorts implicates ion channel genes in the pathogenesis of irritable bowel syndrome. Neurogastroenterol Motil 2018;30: e13358.
[14] Holland D, Frei O, Desikan R, *et al*. Beyond SNP heritability: polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. PLoS Genet 2020;16:e1008612.
[15] Manolio TA, Collins FS, Cox NJ, *et al*. Finding the missing heritability of complex diseases. Nature 2009;461:747–53.
[16] O'Connor LJ. The distribution of common-variant effect sizes. Nat Genet 2021;53:1243–9.
[17] Mathew G, Agha R, Albrecht J, *et al*. STROCSS 2021: Strengthening the reporting of cohort, cross-sectional and case-control studies in surgery. Int J Surg (London, England) 2021;96:106165.
[18] Winkler TW, Day FR, Croteau-Chonka DC, *et al*. Quality control and conduct of genome-wide association meta-analyses. Nat Protoc 2014;9: 1192–212.
[19] Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26:2190–1.
[20] Bulik-Sullivan BK, Loh P-R, Finucane HK, *et al*. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015;47:291–5.
[21] Yin L, Zhang H, Tang Z, *et al*. rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. Genom Proteom Bioinform 2021;19:619–28.
[22] Weeks EM, Ulirsch JC, Cheng NY, *et al*. Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. Nat Genet 2023;55:1267–76.
[23] de Leeuw CA, Mooij JM, Heskes T, *et al*. MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput Biol 2015;11:e1004219.
[24] McLaren W, Gil L, Hunt SE, *et al*. The ensembl variant effect predictor. Genome Biol 2016;17:122.
[25] Watanabe K, Taskesen E, van Bochoven A, *et al*. Functional mapping and annotation of genetic associations with FUMA. Nat Commun 2017;8: 1826.
[26] GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 2020;369:1318–30.
[27] Liberzon A, Subramanian A, Pinchback R, *et al*. Molecular signatures database (MSigDB) 3.0. Bioinformatics 2011;27:1739–40.
[28] Shaw DR. Searching the Mouse Genome Informatics (MGI) resources for information on mouse biology from genotype to phenotype. Curr Protoc Bioinformatics 2016;56:1.7.1–16.
[29] Swan C, Duroudier NP, Campbell E, *et al*. Identifying and testing candidate genetic polymorphisms in the irritable bowel syndrome (IBS): association with TNFSF15 and TNFα. Gut 2013;62:985–94.
[30] Hanscombe KB, Coleman JRI, Traylor M, *et al*. ukbtools: an R package to manage and query UK Biobank data. PLoS One 2019;14:e0214311.
[31] Chang CC, Chow CC, Tellier LC, *et al*. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7.
[32] Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–75.
[33] Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691–2.
[34] Johnstone, IM. Centroid Shrinkage in Statistical Estimation. Journal of the Royal Statistical Society: Series B (Methodological). 1995;57: 76–88.
[35] Bulik-Sullivan B, Finucane HK, Anttila V, *et al*. An atlas of genetic correlations across human diseases and traits. Nat Genet 2015;47: 1236–41.
[36] Skrivankova VW, Richmond RC, Woolf BAR, *et al*. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: The STROBE-MR Statement. JAMA 2021;326: 1614–21.
[37] Hemani G, Zheng J, Elsworth B, *et al*. The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018;7: e34408.
[38] Verbanck M, Chen C-Y, Neale B, *et al*. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. Nat Genet 2018;50: 693–8.
[39] Öhman L, Törnblom H, Simrén M. Crosstalk at the mucosal border: importance of the gut microenvironment in IBS. Nat Rev Gastroenterol Hepatol 2015;12:36–49.
[40] Bennet SMP, Sundin J, Magnusson MK, *et al*. Altered intestinal antibacterial gene expression response profile in irritable bowel syndrome is linked to bacterial composition and immune activation. Neurogastroenterol Motil 2018;30:e13468.
[41] Gribar SC, Richardson WM, Sodhi CP, *et al*. No longer an innocent bystander: epithelial toll-like receptor signaling in the development of mucosal inflammation. Mol Med 2008;14:645–59.
[42] Kao Y-S, Cheng-Chung Wei J. Comment to "Chronic Constipation as a Risk Factor for Colorectal Cancer: Results from a Nationwide, Case-Control Study". Clin Gastroenterol Hepatol 2022;20:2410–1.
[43] Chey WD, Kurlander J, Eswaran S. Irritable bowel syndrome: a clinical review. JAMA 2015;313:949–58.
[44] Wang L-W, Ruan H, Wang B-M, *et al*. Microbiota regulation in constipation and colorectal cancer. World J Gastrointest Oncol 2023;15: 776–86.
[45] Guérin A, Mody R, Fok B, *et al*. Risk of developing colorectal cancer and benign colorectal neoplasm in patients with chronic constipation. Aliment Pharmacol Ther 2014;40:83–92.
[46] Huang W, Tian X, Guan X. The prognosis analysis of RFWD2 inhibiting the expression of ETV1 in colorectal cancer. Transl Cancer Res 2020;9: 508–21.
[47] Siraj AK, Masoodi T, Bu R, *et al*. MED12 is recurrently mutated in Middle Eastern colorectal cancer. Gut 2018;67:663–71.
[48] Ge W, Hu H, Cai W, *et al*. High-risk Stage III colon cancer patients identified by a novel five-gene mutational signature are characterized by upregulation of IL-23A and gut bacterial translocation of the tumor microenvironment. Int J Cancer 2020;146:2027–35.
[49] Wang Y-R, Meng L-B, Su F, *et al*. Insights regarding novel biomarkers and the pathogenesis of primary colorectal carcinoma based on bioinformatic analysis. Comput Biol Chem 2020;85:107229.
[50] Pasman JA, Verweij KJH, Gerring Z, *et al*. GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal influence of schizophrenia. Nat Neurosci 2018;21:1161–70.
[51] Boutwell B, Hinds D. 23andMe Research Team, *et al*. Replication and characterization of CADM2 and MSRA genes on human behavior. Heliyon 2017;3:e00349.
[52] Nagel M, Watanabe K, Stringer S, *et al*. Item-level analyses reveal genetic heterogeneity in neuroticism. Nat Commun 2018;9:905.
[53] Iossifov I, O'Roak BJ, Sanders SJ, *et al*. The contribution of de novo coding mutations to autism spectrum disorder. Nature 2014;515: 216–21.
[54] Moreno V, Alonso MH, Closa A, *et al*. Colon-specific eQTL analysis to inform on functional SNPs. Br J Cancer 2018;119:971–7.
[55] Liu G, Hu Y, Jiang Q. Population difference and disease status affect the association between genetic variants and gene expression. Gastroenterology 2019;157:894–6.
[56] Zhang J, Chen M-J, Zhao G-X, *et al*. Common genetic variants in PRRC2A are associated with both neuromyelitis optica spectrum disorder and multiple sclerosis in Han Chinese population. J Neurol 2021; 268:506–15.
[57] Li C, Chen Y, Chen Y, *et al*. The causal association of irritable bowel syndrome with multiple disease outcomes: a phenome-wide mendelian randomization study. J Clin Med 2023;12:1106.
[58] Gong W, Guo P, Li Y, *et al*. Role of the gut-brain axis in the shared genetic etiology between gastrointestinal tract diseases and psychiatric disorders: a genome-wide pleiotropic analysis. JAMA Psychiatry 2023;80:360–70.

[59] Alemany S, Soler-Artigas M, Cabana-Domínguez J, *et al*. Genome-wide multi-trait analysis of irritable bowel syndrome and related mental conditions identifies 38 new independent variants. J Transl Med 2023;21: 272.
[60] Zamani M, Alizadeh-Tabari S, Zamani V. Systematic review with meta-analysis: the prevalence of anxiety and depression in patients with irritable bowel syndrome. Aliment Pharmacol Ther 2019;50: 132–43.
[61] Lai H-C, Lin H-J, Kao Y-W, *et al*. Irritable bowel syndrome increases the risk of chronic obstructive pulmonary disease: a retrospective cohort study. Sci Rep 2020;10:10008.
[62] Shen T-C, Lin C-L, Wei C-C, *et al*. Bidirectional association between asthma and irritable bowel syndrome: two population-based retrospective cohort studies. PLoS One 2016;11:e0153911.
[63] Freuer D, Linseisen J, Meisinger C. Asthma and the risk of gastrointestinal disorders: a Mendelian randomization study. BMC Med 2022;20:82.