nature portfolio

Corresponding author(s): Terrence Furey

Last updated by author(s): Nov 14, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	CrossMap v0.6.3 Python v3.6.6 plink v1.90b3 Eagle v2.4 R v4.1.3 STAR v2.7.9a verifyBAMID v1.1.3 QTLtools v1.3.1 RUVSeq v1.24.0 kneedle v0.5.0 rtracklayer v1.54.0 qvalue v2.26.0 coloc v5.2.3 locuszoomr v0.2.0 mashr v0.2.79 TOPLD pheatmap v1.0.12 DAVID v2024q1 rrvgo v1.6.0

webr v0.1.5
DESeq2 v1.34.0
edgeR v3.36.0
data.table v1.14.8
tidyverse v2.0.0
ggplot2 v3.5.1
ggsignif v0.6.4
UpSetR v1.4.0
ComplexHeatmap v2.10.0
LDlinkR v1.3.0
AnnotationHub v3.2.2
memoise v2.0.1
METAL v2020-05-05
Code used in this study is available here: https://github.com/nnishiyama/IBD_colon_eQTL

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Processed RNA-seq data are available in the Gene Expression Omnibus (GEO) accession number GSE279302 (https://ww.ncbi.mln.nih.gov/geo/query/acc.cgi?acc=GSE279302). Raw RNA-seq, genotypes, and full eQTL summary statistics are available under restricted access for patient confidentiality. Access can be obtained through dbGaP accession number phs003881.v1.p1. IBD Plexus data are available upon approved application to Crohn's & Colitis Foundation IBD Plexus program (https://www.crohnscolitisfoundation.org/ibd-plexus). Publicly available v8 eQTL summary statistics were downloaded from the GTEx portal (https://gtexportal.org/home/downloads/adult-gtex/overview). Full summary statistics for BarcUVa-seq eQTL were downloaded from the Digital Repository of the University of Barcelona (https://diposit.ub.edu/dspace/handle/2445/172697). Source data are provided with this paper. Lead summary statistics and associated results are provided in the Supplementary Tables/Supplementary Information and Source Data file.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Information about sex was collected at the time of consent and confirmed using genotype array. Approximately 52% of individuals included in this study were female (n = 131; total n = 252). Sex was included as a covariate in the cis-eQTL linear model.
Reporting on race, ethnicity, or other socially relevant groupings	Socially constructed and/or relevant terms including race and ethnicity were not used in this study. Genotypes were used to infer genetic similarity to 1000 Genomes reference superpopulations using principal component analysis (PCA). We included the first four genotype PCs in our cis-eQTL linear model to control for population structure.
Population characteristics	This study focused on a patient cohort diagnosed with inflammatory bowel disease (IBD; Crohn's disease n = 199; ulcerative colitis n = 53). The median age at sample collection was 38 +/- 14.98 years. We identified 84.52% of patients used in this study to have the greatest genetic similarity to the 1000 Genomes reference European superpopulation and 13.10% of patients to have the greatest genetic similarity to the African superpopulation
Recruitment	Patients were recruited at the UNC Multidisciplinary Center for IBD. Written informed consent was obtained from all participants.
Ethics oversight	This study was approved by Human Research Ethics Committee at UNC-Chapel Hill and carried out with accordance to 10-0355 and 11-0359

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size (n=252) was determined based on available sample numbers. Previous colon tissue eQTL studies have successfully detected associations with sample sizes ranging from 48-445 (PMID25766683; PMID30283144; PMID33602935; PMID32913098; PMID33601062).
Data exclusions	Genotype samples with a missingness rate >10% of genotype calls or within two degrees of relatedness were removed. Total gene expression for RNA-seq samples was assessed using principal component analysis (PCA). PCA outliers were removed from downstream analyses. In total, 1 genotype samples and 3 RNA-seq samples were omitted by these criteria.
Replication	cis-eQTL results were compared to external colon eQTL reported by GTEx and BarcUVa-seq (PMID32913098; PMID33601062) in this study. The replication rate was 91.2% for GTEx eQTL and 94.9% for BarcUVa eQTL.
Randomization	No randomization was performed. Covariates including sex, sequencing batch, the first 4 genotype PCs were included in the eQTL linear model. Additional covariates including age, cellular composition, and sequencing quality were corrected for using factors determined by RUVSeq and were also included in the eQTL linear model.
Blinding	Samples were selected based on patient diagnosis at the time of collection. However, investigators were blinded to patient genotypes and other characteristics during library preparation, sequencing, and quality control.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*