nature portfolio

Peer Review File

Integrating genetics and transcriptomics to characterize shared mechanisms in digestive diseases and psychiatric disorders

Corresponding Author: Dr Xiao Chang

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This valuable study integrates genetic and transcriptomic data to investigate shared mechanisms between digestive and psychiatric disorders. The authors employ a diverse methodological approach, effectively combining results to pinpoint relevant genes. While the study yields important findings, particularly through robust transcriptomic analyses, addressing several methodological and presentational aspects could enhance its impact.

1. Regarding the MTAG analyses, several points require clarification:
• The statement in lines 488-489 concerning MTAG needs revision: "This approach enhances the power to detect genetic loci associated with correlated traits by addressing sample overlap and incomplete genetic correlation." While the approach does enhance power by combining data from correlated traits and accounts for sample overlap, the power increase isn't solely due to addressing sample overlap. Please revise for accuracy.
• Was there any sample overlap between the two GWAS? While the manuscript mentions that MTAG can account for sample overlap, it doesn't specify if there was any overlap between the two studies or if the analysis used the no-overlap option. This should be clarified.
• The rationale for applying MTAG solely to IBS and GERD, and not to other traits with similar genetic correlation patterns, lacks clarity and requires further explanation.

2. The methods section requires additional detail to ensure reproducibility:
• Citations should be integrated throughout to reference specific approaches. For example, Finucane et al. (Nature Genetics 2015) should be cited when describing S-LDSC, and the relevant MTAG paper should be referenced in its corresponding section.
• In line 473, when mentioning "hierarchical clustering techniques," the specific technique employed and the software used should be specified.
• In line 525, where "seven brain tissues and seven digestive tract tissues" are mentioned, please list the specific tissues included and explain the selection criteria.

3. In Table S2, there appears to be an error where some p-values exceed their corresponding FDR q-values. For instance, this occurs for IBS and anorexia nervosa (0.29 and 4.8e-3, respectively). Please carefully review the table for accuracy and ensure that p-values are always smaller than or equal to their respective q-values.

4. Minor Comments:
• In Table S36, the header misspells "Module."
• Line 213 incorrectly references supplementary figures; they should be S9 and S10. These colocalization plots also require textual descriptions.
• In the LD score formula, "rg" needs definition.
• Line 537's "This correlation matrix was transformation" should be corrected to "was transformed."
• Figure 1 and S11 require an explanation of the asterisk (*) symbol.

Reviewer #2

(Remarks to the Author)
In this work, the authors aim to explore the genetic connections between digestive and psychiatric disorders through the integration of genomic and transcriptomic data. The approach is interesting overall, but several minor revisions are

necessary to enhance the manuscript's clarity, strengthen the methodology, and improve the overall presentation to meet publishable standards

---

1.The abstract should begin with a brief background that highlights the significance of investigating the links between psychiatric and gastrointestinal disorders, providing essential context for the study.
2.It should also include a concise summary of the specific methods or approaches used to integrate genetic and transcriptomic data, ensuring clarity and transparency in the research process.
3.Additionally, the abstract should interpret the findings, explaining how the results enhance our understanding of the relationship between these disorders. And it should mention potential applications or future research directions, emphasizing the broader impact of the study on the field.
4. In the introduction, please simplify those complex sentences to improve readability. For example, the term "these" in line 79 is unclear and needs clarification. Similarly, the phrase "The expansion of sample sizes..." should be rephrased for better understanding. Such clarifications are needed throughout the introduction.
5, There are a few terms that were misused. For instance, the word "significant" in line 64 should be reserved for results that are statistically validated, maintaining proper terminology usage.
6. Only one Mendelian Randomization (MR) study is referenced. A broader review of existing research on the relationship between psychiatric disorders (e.g., depression) and gastrointestinal disorders, using MR or other genetic approaches, is necessary. The authors should summarize the current research landscape, cite relevant studies, identify gaps, and explain how this study addresses them.
7. In the methods section, standardize the descriptions of the GWAS datasets used for psychiatric disorders and digestive diseases. Currently, some sections, such as GERD, provide more detail than others, like NE. Ensure consistent formatting and detail across all datasets for clarity.
8. The full name of a few abbreviations and acronyms were not mentioned upon their first place (e.g., "LDSC" is used without explanation). Reintroduce abbreviations if they reappear after a long gap to aid reader comprehension.
9. The description of the MAGMA method lacks clarity. The authors should include specific details about its application: Was gene analysis conducted for each trait separately, or were common genes identified across traits? The same level of detail should be applied to the PPI network analysis.
10. The section for study limitation is missing.

Reviewer #3

(Remarks to the Author)
This article explores the connection between mental disorders and digestive system diseases by combining genetic and transcriptomic data. It highlights significant genetic correlations between conditions like irritable bowel syndrome (IBS), gastroesophageal reflux disease (GERD), depression (DEP), and neuroticism (NE). Through multi-trait GWAS analysis, pleiotropic loci linked to both IBS and GERD are found, especially those related to mental health aspects. Moreover, a gene co-expression network analysis reveals neuro-pathway-related gene modules enriched in digestive tissues. These findings provide new understanding of the shared biological mechanisms between mental and digestive disorders.

Here are some revision suggestions for the study:
1.Gene Modules (Lines 243-244): There is a notable difference in the number of co-expression modules across different tissues, ranging from 4 to 17. It is recommended that the authors explain the potential reasons for this variation and its impact on the interpretation of the results.
2.Protein-Protein Interaction (PPI) Analysis (Lines 561-569): The article conducts PPI analysis on overlapping genes within co-expression modules of different tissues. It is suggested that the authors provide a detailed explanation of the methods and parameters used to construct the PPI network and how the PPI results are interpreted.
3.Module Preservation Analysis (Lines 546-558): The study uses the Zsummary metric to assess module preservation. The authors are encouraged to explain the rationale for selecting this metric and discuss the biological significance of different thresholds (e.g., >20, >10, 2-10, <2).
4. Sample overlap: It is recommended that the authors clarify in the methods section whether potential sample overlap between the two GWAS was considered in the MTAG analysis and how this was addressed.
5.Research Limitations: It is advised that the authors enrich the discussion section with a comprehensive explanation of the study's limitations, including but not limited to sample representativeness, potential biases in data sources, and the generalizability of the results.
6. In the "Abstract" section, the sentence "our network analysis suggests BSN, CELF4, and NRXN1 as central players in the regulation of gut-brain axis in digestive diseases." should include the definite article "the" before "gut-brain axis."
7. In the "Introduction" section, the sentence "to address critical questions such as: 1) Is the presence of GBA associated genes a primary and determining factor influencing the development of digestive tract diseases?" should change "GBA associated genes" to "GBA-associated genes" for consistency.
8. Clearly define all abbreviations and terms upon their first appearance to ensure readers can easily understand.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

After incorporating the reviewers' comments and suggestions, the authors have successfully improved the manuscript, making it suitable for acceptance and publication.

Reviewer #2

(Remarks to the Author)
The authors have properly addressed my concerns.

Reviewer #3

(Remarks to the Author)
The authors adequately addressed my previous comments.

Reply to Reviewers' comments

We express our gratitude to the editor and the reviewers for their meticulous evaluation of our manuscript and for their valuable suggestions. In response to the feedback provided, we have made thorough revisions to the manuscript to comprehensively address each of the reviewers' comments. Below is a detailed point-by-point response:

Reviewer #1:

This valuable study integrates genetic and transcriptomic data to investigate shared mechanisms between digestive and psychiatric disorders. The authors employ a diverse methodological approach, effectively combining results to pinpoint relevant genes. While the study yields important findings, particularly through robust transcriptomic analyses, addressing several methodological and presentational aspects could enhance its impact.

1. Regarding the MTAG analyses, several points require clarification:

• The statement in lines 488-489 concerning MTAG needs revision: "This approach enhances the power to detect genetic loci associated with correlated traits by addressing sample overlap and incomplete genetic correlation." While the approach does enhance power by combining data from correlated traits and accounts for sample overlap, the power increase isn't solely due to addressing sample overlap. Please revise for accuracy.

Reply: Thank you for pointing this out. We have revised the statement in the Methods section to improve accuracy: “This approach enhances the ability to detect loci from related traits by jointly analyzing GWAS summary statistics. Compared to traditional inverse-variance weighted meta-analysis, it also accounts for sample overlap and incomplete genetic correlation.” (Page 26, Lines 522-525).

• Was there any sample overlap between the two GWAS? While the manuscript mentions that

MTAG can account for sample overlap, it doesn't specify if there was any overlap between the two studies or if the analysis used the no-overlap option. This should be clarified.

Reply: We appreciate the reviewer's insightful comments. In the present study, the GWAS data used for MTAG were drawn from distinct research cohorts, and theoretically, sample overlap among these cohorts should not occur. However, it is crucial to acknowledge that all these studies rely on large-scale populations, and definitively ruling out potential overlaps is challenging, especially when utilizing publicly available datasets. Considering this, we opted to use the default settings in the MTAG analysis without explicitly enabling the "no-overlap option," which is designed to assume the absence of overlap between cohorts. By doing so, we allowed MTAG to account for potential sample overlap. We have clarified this point in the Methods section and addressed the potential impact of any sample overlap as a limitation in the Discussion section of the revised manuscript. (Page 21, Lines 428-431).

• *The rationale for applying MTAG solely to IBS and GERD, and not to other traits with similar genetic correlation patterns, lacks clarity and requires further explanation.*

Reply: We thank the reviewer for raising this important point. The decision to apply MTAG solely to IBS and GERD, rather than to other traits with similar genetic correlation patterns, was driven by the significant genetic correlations these two digestive disorders share with psychiatric conditions. IBS and GERD showed particularly strong associations with psychiatric traits such as depression, neuroticism, and major depressive disorder, as highlighted by linkage disequilibrium score regression (LDSC) analyses, which revealed substantial genetic overlap between these disorders and psychiatric conditions. This focus is further supported by the identification of numerous novel loci near genes implicated in neurological pathways, which are relevant to both digestive and psychiatric traits.

In contrast, MTAG analyses of other digestive disorders with similar genetic correlation patterns yielded far fewer significant loci, particularly those linked to psychiatric traits. These results, which provide additional context, have been included in the supplementary materials

(Supplementary Tables 11–32). We have clarified this rationale in the revised manuscript to ensure greater transparency. (Page 12, Line 225-230)

2. The methods section requires additional detail to ensure reproducibility:

• Citations should be integrated throughout to reference specific approaches. For example, Finucane et al. (Nature Genetics 2015) should be cited when describing S-LDSC, and the relevant MTAG paper should be referenced in its corresponding section.

Reply: Thank you for the suggestion. We have included the appropriate references in the S-LDSC and MTAG sections for clarity and reproducibility (Page 24, Line 495, References 49, 50; Page 26, Line 522, Reference 14).

• In line 473, when mentioning "hierarchical clustering techniques," the specific technique employed and the software used should be specified.

Reply: We appreciate the reviewer's insightful comments. We made modifications in the S-LDSC methods section: “The enrichment values specific to each annotation are converted into a color scale and visualized using hierarchical clustering techniques in the ComplexHeatmap package in R 4.2.3. Specifically, hierarchical clustering is performed using Euclidean distance as the distance metric and complete linkage as the clustering method, which are the default settings in the ComplexHeatmap package (internally relying on the hclust function from base R).” (Page25, Lines 502-507)

• In line 525, where "seven brain tissues and seven digestive tract tissues" are mentioned, please list the specific tissues included and explain the selection criteria.

Reply: Thank you for the suggestion. The specific tissues mentioned are detailed in Supplementary Table S34. The seven brain tissues include: Brain – Cortex, Brain – Caudate (basal ganglia), Brain – Hypothalamus, Brain – Amygdala, Brain – Hippocampus, Brain – Cerebellar Hemisphere, and Brain – Spinal Cord (cervical c-1). The seven digestive tract tissues include:

Esophagus – Mucosa, Esophagus – Muscularis, Esophagus – Gastroesophageal Junction, Stomach, Small Intestine – Terminal Ileum, Colon – Transverse, and Colon – Sigmoid.

The selection criteria were guided by existing literature. The seven brain tissues were chosen due to their potential involvement in psychiatric disorders analyzed in this study, such as depression (DEP), major depressive disorder (MDD), and neuroticism (NE). Similarly, the seven digestive tract tissues were selected for their relevance to the development and progression of GERD and IBS. We have added this clarification to the revised manuscript to ensure transparency (Page 12-13, Lines 245-249).

3. In Table S2, there appears to be an error where some p-values exceed their corresponding FDR q-values. For instance, this occurs for IBS and anorexia nervosa (0.29 and 4.8e-3, respectively). Please carefully review the table for accuracy and ensure that p-values are always smaller than or equal to their respective q-values.

Reply: We greatly appreciate the reviewer's feedback and apologize for the errors identified in Table S2. These discrepancies were due to inaccuracies during data organization. We have corrected the issues in the revised version of Table S2 and conducted a thorough review to ensure that all p-values are smaller than or equal to their respective q-values. Thank you for bringing this to our attention.

4. Minor Comments:

• In Table S36, the header misspells "Module."

Reply: We thank the reviewer for pointing out this error. The spelling mistake in the header has been corrected in the revised Table S36.

• Line 213 incorrectly references supplementary figures; they should be S9 and S10. These colocalization plots also require textual descriptions.

Reply: We appreciate the reviewer's detailed comments. The description errors related to Supplementary Figures 9 and 10 have been corrected (Page 11, Lines 216). Additionally, we have added a detailed description of the colocalization analysis in the supplementary materials for clarity and completeness.

• *In the LD score formula, "rg" needs definition.*

Reply: We thank the reviewer for their thorough examination. Genetic correlation (R_g) has been defined in the Methods section under LDSC (Page 24, Lines 484-485) in the revised manuscript.

• *Line 537's "This correlation matrix was transformation" should be corrected to "was transformed."*

Reply: We thank the reviewer for their thorough review. The error has been corrected, and "This correlation matrix was transformation" has been updated to "was transformed" (Page 28, Line 576) in the revised manuscript.

• *Figure 1 and S11 require an explanation of the asterisk (*) symbol.*

Reply: We thank the reviewer for their careful review. The explanation for the asterisk symbol has been added to the figure legends of Figures 1 and S11: '*FDR < 0.05, **FDR < 0.01, ***FDR < 0.001.'

Reviewer #2:

In this work, the authors aim to explore the genetic connections between digestive and psychiatric disorders through the integration of genomic and transcriptomic data. The approach is interesting overall, but several minor revisions are necessary to enhance the manuscript's clarity, strengthen the methodology, and improve the overall presentation to meet publishable standards

1.The abstract should begin with a brief background that highlights the significance of investigating the links between psychiatric and gastrointestinal disorders, providing essential context for the study.

Reply: We appreciate the reviewer for pointing this out, and we have added a brief rationale at the beginning of the abstract (Page 3, Lines 38).

2.It should also include a concise summary of the specific methods or approaches used to integrate genetic and transcriptomic data, ensuring clarity and transparency in the research process.

Reply: We appreciate the reviewer's valuable suggestions. We have included the main analytical methods in the abstract (Page 3, Lines 39-41).

3.Additionally, the abstract should interpret the findings, explaining how the results enhance our understanding of the relationship between these disorders. And it should mention potential applications or future research directions, emphasizing the broader impact of the study on the field.

Reply: We thank the reviewer for the valuable suggestion. We have added the implications of the research findings at the end of the abstract (Page 3, Lines 50-52).

4. In the introduction, please simplify those complex sentences to improve readability. For example, the term "these" in line 79 is unclear and needs clarification. Similarly, the phrase "The expansion of sample sizes..." should be rephrased for better understanding. Such clarifications are needed throughout the introduction.

Reply: We appreciate the reviewer for the constructive suggestion. We have optimized complex sentences for clarity. Additionally, we have clarified what 'these' and 'the expansion of the sample size...' refer to (Page 4, Lines 68-70; Page 4-5, Lines 73-80).

5, There are a few terms that were misused. For instance, the word "significant" in line 64 should be reserved for results that are statistically validated, maintaining proper terminology usage.

Reply: We appreciate the reviewer's comment. The wording has been revised to ensure accurate terminology usage (Page 4, Lines 60).

6. Only one Mendelian Randomization (MR) study is referenced. A broader review of existing research on the relationship between psychiatric disorders (e.g., depression) and gastrointestinal disorders, using MR or other genetic approaches, is necessary. The authors should summarize the current research landscape, cite relevant studies, identify gaps, and explain how this study addresses them.

Reply: We appreciate the reviewer's insightful suggestions. Building on the modifications made in the fourth point, we have expanded the background to include a broader review of existing GWAS and MR studies on the gut-brain axis, which primarily provide a genomic overview (Page 4-5, Lines 73-85). Relevant studies have been cited to identify research gaps, and the final paragraph now explains how our study addresses these gaps (Page 5-6, Lines 95-103).

7. In the methods section, standardize the descriptions of the GWAS datasets used for psychiatric disorders and digestive diseases. Currently, some sections, such as GERD, provide more detail than others, like NE. Ensure consistent formatting and detail across all datasets for clarity.

Reply: We sincerely appreciate the reviewer's valuable suggestion. For the sake of consistency, we have added detailed information about the GWAS data on mental disorders in the methods section (Page 23-24, Lines 469-476).

8. The full name of a few abbreviations and acronyms were not mentioned upon their first place (e.g., "LDSC" is used without explanation). Reintroduce abbreviations if they reappear after a long gap to aid reader comprehension.

Reply: We appreciate the reviewer's valuable comments. The manuscript has been thoroughly reviewed to ensure that all abbreviations are clearly defined upon their first appearance and reintroduced where necessary to maintain reader comprehension.

9. The description of the MAGMA method lacks clarity. The authors should include specific details about its application: Was gene analysis conducted for each trait separately, or were common genes identified across traits? The same level of detail should be applied to the PPI network analysis.

Reply: Thank you for the reviewer's valuable comments. We have provided a more detailed description of MAGMA (Page 27, Lines 551-555) and PPI network analysis (Page 31, Lines 620-630).

10. The section for study limitation is missing.

Reply: We sincerely appreciate the reviewer's suggestion. A discussion of the study's limitations has been added at the end of the discussion section to address this point (Page 21, Lines 425-433).

Reviewer #3:

This article explores the connection between mental disorders and digestive system diseases by combining genetic and transcriptomic data. It highlights significant genetic correlations between conditions like irritable bowel syndrome (IBS), gastroesophageal reflux disease (GERD), depression (DEP), and neuroticism (NE). Through multi-trait GWAS analysis, pleiotropic loci linked to both IBS and GERD are found, especially those related to mental health aspects. Moreover, a gene co-expression network analysis reveals neuro-pathway-related gene modules enriched in digestive tissues. These findings provide new understanding of the shared biological mechanisms between mental and digestive disorders.

Here are some revision suggestions for the study:

1. Gene Modules (Lines 243-244): There is a notable difference in the number of co-expression modules across different tissues, ranging from 4 to 17. It is recommended that the authors explain the potential reasons for this variation and its impact on the interpretation of the results.

Reply: Thank you for this insightful comment. The variation in the number of co-expression modules across tissues likely reflects differences in their biological complexity and functional roles. For instance, tissues like the sigmoid colon, which show 17 modules, may have more diverse regulatory networks compared to simpler tissues like the stomach, which has only 4 modules. Additionally, the number of samples analyzed and the diversity of tissue-specific gene expression profiles can also influence the number of detected modules, with larger sample sizes often increasing statistical power to identify more modules. This variation emphasizes the unique regulatory demands of each tissue and highlights the importance of considering tissue-specific gene networks in the analysis. We have added this explanation to the revised manuscript (Pages 18, Lines 354-358).

2. Protein-Protein Interaction (PPI) Analysis (Lines 561-569): The article conducts PPI analysis on overlapping genes within co-expression modules of different tissues. It is suggested that the authors provide a detailed explanation of the methods and parameters used to construct the PPI network and how the PPI results are interpreted.

Reply: Thank you for the valuable comment. We have provided a more detailed description of the PPI network analysis (Page 30, Lines 620-630).

3. Module Preservation Analysis (Lines 546-558): The study uses the Zsummary metric to assess module preservation. The authors are encouraged to explain the rationale for selecting this metric and discuss the biological significance of different thresholds (e.g., >20, >10, 2-10, <2).

Reply: Thank you for the valuable comment. We have provided a detailed description of the

$Z_{summary}$ metric in the results and methods sections (Pages 29-30, Lines 595-617).

4. Sample overlap: It is recommended that the authors clarify in the methods section whether potential sample overlap between the two GWAS was considered in the MTAG analysis and how this was addressed.

Reply: Thank you for raising this important point. The GWAS datasets used in the MTAG analysis were obtained from separate research cohorts, making sample overlap unlikely. However, given the large-scale nature of these datasets, the possibility of minor overlaps cannot be completely excluded. To account for this, we used the default MTAG settings, which inherently manage potential sample overlap without requiring additional adjustments. This approach, along with its potential limitations, has been clarified in the discussion section of the revised manuscript (Pages 21, Lines 428-431).

5. Research Limitations: It is advised that the authors enrich the discussion section with a comprehensive explanation of the study's limitations, including but not limited to sample representativeness, potential biases in data sources, and the generalizability of the results.

Reply: Thank you very much for the valuable suggestion. We have added a detailed discussion of the study's limitations, addressing sample representativeness, potential biases in data sources, and the generalizability of the results, at the end of the discussion section (Pages 21, Lines 425-433).

6. In the "Abstract" section, the sentence "our network analysis suggests BSN, CELF4, and NRXN1 as central players in the regulation of gut-brain axis in digestive diseases." should include the definite article "the" before "gut-brain axis."

Reply: Thank you very much for the careful review. We have added "the" before "gut-brain axis" in the Abstract section (Page 3, Line 50).

7. In the "Introduction" section, the sentence "to address critical questions such as: 1) Is the presence of GBA associated genes a primary and determining factor influencing the development of digestive tract diseases?" should change "GBA associated genes" to "GBA-associated genes" for consistency.

Reply: Thank you very much for the careful review. We have updated "GBA associated genes" to "GBA-associated genes" for consistency (Page 5, Line 86).

8. Clearly define all abbreviations and terms upon their first appearance to ensure readers can easily understand.

Reply: Thank you for the valuable comment. We have thoroughly reviewed the manuscript to ensure that all abbreviations and terms are clearly defined upon their first appearance for improved readability and comprehension.