nature portfolio

Peer Review File

Cell-type-aware Transcriptome-wide Association Studies Identify 91 Independent Risk Genes for Alzheimer's Disease Dementia

Corresponding Author: Dr Jingjing Yang

This manuscript has been previously submitted at another journal. This document only contains information relating to versions considered at Communications Biology.

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Liu et al. present an interesting study, leveraging state of the art omics data and analytical tools to enable integration with genetics, to identify likely causal Alzheimer's disease genes in brain bulk tissue and cell-type specific contexts. The manuscript is relevant to the journal and I am generally enthusiastic about its impact. However, there are several points where I believe analysis and writing needs additional clarification.

Major
- The results section is too long. The first section that present a methods overview is out of place. Even if the authors feel that presenting some details on methods in the results section is relevant, this should be scattered throughout the results at relevant subsections and kept minimal. Descriptions of findings per cell-type are extensive and repetitive; this could easily be condensed. Overall, the current presentation of the results section decreases readability and detracts from the impact.
- The authors use the novel TWAS method TIGAR/DPR in addition to the more common Fusion and PrediXcan methods. It should be motivated in more detail why this is done and what is the diRerence/added value of TIGAR. It was very striking that TIGAR produced ~13k gene imputation models for diRerent cell types with cTWAS, while only ~4-5k were observed for the other 2 methods, yet, this diRerence is not observed for PWAS where there are ~4k protein across all methods. No mention is made of this in the manuscript. This raises concerns about technical issues or false positive findings with TIGAR in the cell-specific analyses.
- It is not clear why diRerent statistical thresholds are used for cell-specific TWAS versus PWAS. It is also not explained why the threshold is set at P<2.5e-6 for a subset of the analyses. This raises questions about the validity of the statistical criteria and thus needs careful clarification.
- The fine mapping method, GIFT, needs to be explained in more detail so it becomes easier for readers to understand the related results. Some figures are potentially confusing, e.g. figure 5, top middle panel: Why is LIMS2 prioritized with GIFT, but not the other 2 TWAS hits on the left (that are of much higher significance)? Is this because GIFT could not resolve them? This needs to be clarified more. Also, in figure 4, why does chr17 not have a red dot prioritized by GIFT?
- I have doubt that the APOC2 finding is true. Firstly, it seems the manuscript states that APOC2 encodes a protective apoE isoform, but APOC2 and APOE are diRerent genes; this requires correction. More critically, the APOE locus is very complex and marked by strong linkage making it easy to obtain false positive variant or gene

associations. Without a more careful and granular assessment of this finding and the locus, the validity of this finding will remain a concern. Importantly also, the authors use Bellenguez et al. 2022 summary statistics, but those statistics are flawed at the APOE locus because P-values have been capped of at P<1e-300, which results in the APOE*4 variant, rs429358, and potential other variants being cut out of the summary statistics.
- The authors have compared their current results with their previous work (reference 20, https://doi.org/10.1186/s13195-025-01774-y). In the supplementary of this previous published paper, there were 113 significant bulk TWAS-O findings, which is inconsistent with the “Table S1. Bulk TWAS-O” of the current manuscript.
- Adding captions for each Supplementary Tables will be helpful. In addition, for Tables S3–S8, please consider adding extra columns indicating whether each gene is validated by bulk TWAS-O or bulk PWAS-O. The additional columns could explicitly list the corresponding gene, chromosome, and bp position from the bulk analyses when overlap is detected.
- Line 142-145, The description of bulk TWAS-O and bulk PWAS-O results is confusing. Please clarify explicitly that bulk TWAS-O comes from previously published work (ref. 20), whereas bulk PWAS-O was newly performed in this study. In addition, it would be helpful to indicate in the schematic (Figure 1) how these two bulk datasets were used for comparison.

Minor
- It’s great to see QQ plots that indicate the inflation of TWAS/PWAS analyses, but it would be important to also report/add Genomic inflation factors so the readers can have a clearer sense of inflation.
- Abbreviations should be checked for consistency. For example, “Alzheimer’s disease“ is written out multiple times after the abbreviation has already been introduced.
- Line 88, “cix-xQTL” should be “cis-xQTL”.
- In Table 1, the “2” in R^2 should be formatted as a superscript.
- Fig .2, the y-axis label “-log10(q-value)” should be “-log10(p-value)”
- Update your published paper (reference 20).
- Line 264-265, “We identified 7 genes ……” is inconsistent with “ARL4A, RAB8B, HEXIM1, ZNF225, SNRPD2, and FBXO46” (6 genes).
- Many red dots were not labeled in Figure 2B-2F.
- Line 544-550, please present the specific setting of generating PPI network in STRING.
- In the abstract and main manuscript, please clarify that the 223 TWAS risk genes are unique genes derived from 350 significant gene–cell-type associations, to avoid ambiguity in the presentation.
- Line 145-147, it will be helpful to define the validation threshold of bulk TWAS-O.
- Line 301, Line 848, and Line 857, it’s not appropriate to use “physical”, since many PPI links were not physical interaction.
- Line 389, what is “in this previous study”?
- Figure 4 was not mentioned in the main manuscript and is identical to Figure S13.
- There are several inconsistencies need to be updated: Line 213 (20 & 16 flipped); Line 179, “19” inconsistent with the counts in Table S3.
- In Figure S19, the labels bulk and proteome are potentially confusing. It would be clearer to explicitly indicate them as bulk TWAS-O and bulk PWAS-O, respectively, to avoid ambiguity.

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Biology initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Reviewer #3

(Remarks to the Author)
This study by Liu et al. reports an omnibus methodology to identify cell type–specific risk genes. The authors utilised single-nuclei transcriptomic data to identify cell-type specific candidate genes and validated several associations through integration with bulk transcriptomics and proteomics data. They further employed PPI network data and pathway analysis to link these associations to functionally relevant biological pathways. Importantly, the study leverages one of the largest datasets currently available on Alzheimer’s disease (AD), reporting 11 novel associations. To strengthen the manuscript, I would encourage the authors to highlight not only the findings themselves but also their potential impact. While the methodology is improved, there are several limitations that should be addressed before publication.
1. A number of the reported cell type–specific risk genes appear to arise from non-specific expression. snRNAseq data is

inherently prone to expression dropout and mis-annotation of poor quality cells. For example, one astrocyte-specific risk gene reported, CYP19A1, shows extremely low expression in astrocyte according to Human Protein Atlas data. Similarly, one of the genes validated through PWAS (C1QL1) is known to be myeloid cell-specific. Furthermore, GFAP, which is a canonical astrocyte marker, was reported here as a risk gene for oligodendrocytes. These examples suggest that the current analysis may have included spurious associations. I therefore recommend applying a minimum expression threshold or additional quality filters when deriving cell-specific associations from snRNA-seq data.
2. In a recently published study (Haglund et al., https://www.nature.com/articles/s41588-024-02050-9) , a different cohort of snRNA-seq data was used to map cell type–specific eQTLs. Comparing the current findings against this independent dataset would strengthen the confidence in the identified targets.
3. The pragmatic choice of CV R2>0.5% as a threshold is understandable given sample size limitations. However, such a low threshold also risks including weakly predicted genes. I recommend testing higher cutoffs (e.g. 1% or 5%) and reporting how sensitive the results are to this choice, as this would provide readers with a clearer picture of the robustness of the findings.
4. As per my understanding, none of the xWAS method directly tests for colocalization. Therefore, I also recommend performing colocalization analyses to help distinguish true causal relationships from signals driven by linkage disequilibrium. This would substantially increase the mechanistic interpretability of the results.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the comments on behalf of both of us.

Reviewer #2

(Remarks to the Author)
I thank the authors for their detailed and thoughtful responses. They have made substantial efforts to address the reviewers' comments, and most concerns have been adequately resolved. Only one major issue remains that should be addressed prior to publication, while two additional points would benefit from minor clarification to further improve the manuscript.

Major (strongly recommended to be addressed before publication)
Without explicitly conditioning on APOE*4 and APOE*2 allele dosage, reliable gene-level inference at the APOE locus is not possible due to extreme linkage disequilibrium and dominant APOE-driven effects. Consequently, the conclusions regarding APOC2 are not sufficiently supported.
The authors should therefore either adjust for apoe4 and apoe2 dosage for analyses within the APOE locus or completely remove the locus and related discussion from the manuscript.

Minor (partially addressed; suggested minor revisions)
Line 539–540. The revised wording remains potentially misleading. The statement that "DPR yields substantially more valid expression prediction models, particularly for genes with relatively low expression heritability" could be interpreted as implying improved predictive reliability. In practice, DPR primarily increases model permissiveness and the number of fitted models, rather than predictive accuracy per se. A slight rephrasing to clarify this distinction would further improve accuracy and clarity.

Line 372–377. This comparison has been partially clarified, but it may still be misleading. The increased number of significant findings is likely driven by the substantially larger set of genes tested using TIGAR/DPR-based models, rather than by the use of a more stringent significance threshold itself. Explicitly stating this would help avoid overinterpretation and further strengthen the presentation.

Reviewer #3

(Remarks to the Author)
Thanks to authors for addressing the concerns appropriately. The PPI network and pathway figures (figure 4b,c and 5 are distorted and not aesthetically pleasant. I recommend to properly format the figures suitable for publication.

Responses to Reviewers

We thank the editors and reviewers for their thoughtful and constructive comments and for recognizing the importance of our cell-type–aware TWAS analyses for studying Alzheimer's disease (AD). We have thoroughly revised this manuscript to address all Reviewers' comments, which greatly improved this manuscript. Below, we respond point-by-point responses addressing all reviewers' comments.

Reviewer #1 (Remarks to the Author):

Liu et al. present an interesting study, leveraging state of the art omics data and analytical tools to enable integration with genetics, to identify likely causal Alzheimer's disease genes in brain bulk tissue and cell-type specific contexts. The manuscript is relevant to the journal and I am generally enthusiastic about its impact. However, there are several points where I believe analysis and writing needs additional clarification.

Major Comments

1. - The results section is too long. The first section that present a methods overview is out of place. Even if the authors feel that presenting some details on methods in the results section is relevant, this should be scattered throughout the results at relevant subsections and kept minimal. Descriptions of findings per cell-type are extensive and repetitive; this could easily be condensed. Overall, the current presentation of the results section decreases readability and detracts from the impact.

Response: We thank the reviewer for this helpful suggestion. In this revision, we have substantially shortened the Results section by:

> (i) Condensing the "Methods overview" subsection to include only the essential analytical procedure in the revised first Results subsection "Overview of the analytical procedure", in pages 4-5, lines 90-130). We feel like these descriptions would help readers to understand our study flow and the follow up results. Details about the xWAS-O pipeline and ROS/MAP omics data are integrated into the Method section in the end.
>
> (ii) Condensing repetitive cell-type–aware TWAS-O results descriptions, with emphasis on the number of total identified TWAS-O risk genes, fine-mapped independent signals, and novel findings (subsection "Cell-type–aware TWAS-O results", lines 132–198 on pages 6-8).

These changes reduce redundancy, improve readability, and help focus the Results on the most important biological and methodological insights.

2. - The authors use the novel TWAS method TIGAR/DPR in addition to the more common Fusion and PrediXcan methods. It should be motivated in more detail why this is done and what is the difference/added value of TIGAR. It was very striking that TIGAR produced ~13k gene imputation models for different cell types, while only ~4-5k were

observed for the other 2 methods, yet, this difference is not observed for PWAS where there are ~4k protein across all methods. No mention is made of this in the manuscript. This raises concerns about technical issues or false positive findings with TIGAR in the cell-specific analyses.

Response: We thank the reviewer for raising this important point. In this revision, we have added detailed description of the xWAS-O method in the Methods subsection "xWAS-O analytical framework" on pages 18-20. We described the Aggregated Cauchy Association Test which combines the test p-values by TIGAR/DPR, PrediXcan/EN, and FUSION to achieve optimal power while calibrating type I errors shown by comprehensive simulation and real studies in our previous publications (see lines 514-526 on pages 19-20; Nagpal et. al, AJHG 2019, PMID: 31230719; Parrish et. al. HGGA 2021, PMID: 35047855).

In the revised Results section, lines 137-143 on page 6, we now described the reason about why TIGAR produced ~13K gene imputation models versus ~4K by PrediXcan/EN and ~5K by FUSION, cited as follows:

> *"The substantially larger number of valid expression models from TIGAR/DPR is only because the nonparametric Bayesian DPR model best models the genetic architecture underlying most gene expression quantitative traits. The PrediXcan/EN method obtained much less valid prediction models for expression traits, which is mainly due to the limitation of EN method with equal proportion of L1 and L2 penalties that can only select cis-eQTL with relatively large effect sizes."*

In the revised Results section, lines 206-211 on page 8, we now described the reason about why all method obtained similar numbers of imputation models for proteins, cited as follows:

> *"Because a sparse cis-genetic architecture with low heritability is observed for most quantitative protein traits, both TIGAR/DPR and FUSION/BestModel obtained similar number of valid prediction models. The much smaller number of valid protein abundance imputation models obtained by PrediXcan/EN is still due to the limitation of the Elastic-Net method with equal proportion of L1 and L2 penalties that can only select cis-eQTL with relatively large effect sizes."*

To clarify these points, we also added a subsection "Motivation for including TIGAR/DPR in xWAS-O" in page 20 in the Method, explaining the DPR model and why it yields more valid gene expression imputation models.

3. - It is not clear why different statistical thresholds are used for cell-specific TWAS versus PWAS. It is also not explained why the threshold is set at P<2.5e-6 for a subset of the analyses. This raises questions about the validity of the statistical criteria and thus needs careful clarification.

Response: We thank the reviewer for the comment. In this revision, we have added detailed explanations in the Results subsection "Overview of the analytical procedure", lines 103-110 on page 5. See cited texts as follows:

> *"For each cell type, genes with omnibus TWAS p-values < 2.5E-06 (obtained by ACAT[25]) are defined as significant TWAS-O findings, which reflects a conservative genome-wide significance level corresponding to Bonferroni correction for ~20K independent genome-wide genes. Since PWAS-O tests a smaller set of protein coding genes (~5K) and exhibited mild global inflation with genomic control factor $\lambda = 1.47$ (**Fig. S11**), we first adjusted the PWAS-O p-values*

according to the genomic control factor[26], and then derived q-values to control for false discovery rates (FDRs) by the Benjamini-Hochberg procedure[27]. Genes with PWAS-O q-values<0.05 are defined as significant PWAS-O findings."

We clarified that the genome-wide significance threshold of p-value < 2.5E-06 is determined by Bonferroni correction of 20K independent test genes, which is more conservative than the FDR correction by the Benjamini-Hochberg procedure.

Because the total number of tested protein coding genes is only ~5K, much less than the assumed 20K independent tests to use the significance level of 2.5E-06. Thus, we used the FDR (<0.05) derived from PWAS-O p-values after correcting for the genomic control factor to identify significant genes. This FDR-based approach was also employed in our previous PWAS-O paper (T Hu et al. Am J Hum Genet. 2024, PMID: 39079537), which demonstrated calibrated type I error in comprehensive NULL simulation studies.

4. - The fine mapping method, GIFT, needs to be explained in more detail so it becomes easier for readers to understand the related results. Some figures are potentially confusing, e.g. figure 5, top middle panel: Why is LIMS2 prioritized with GIFT, but not the other 2 TWAS hits on the left (that are of much higher significance)? Is this because GIFT could not resolve them? This needs to be clarified more. Also, in figure 4, why does chr17 not have a red dot prioritized by GIFT?

Response: We thank the reviewer for this comment. In this revision, we briefly described how we employed GIFT for fine-mapping in this study in the Results subsection "Overview of the analytical procedure" on lines 111-116 on page 5. We also included more details about the GIFT method in the Methods subsection "Fine-map xWAS-O risk genes by GIFT", lines 550 – 589 on page 20-22).

Specifically, GIFT performs joint, conditional modeling across all genes within a genomic region, accounting for both (i) LD among cis-SNPs and (ii) correlations among genetically regulated gene expression/protein abundance. Since GIFT fine-maps genes using a joint model of multiple genes, the gene with the most significant marginal TWAS/PWAS p-value may not always be prioritized as the top significant gene by GIFT. We clarified this in the Results in lines 151-153 on page 6, and Methods description in lines 580-582 on page 21. This explains why LIMS2 in **Fig. 3** (Fig. 4 in initial submission) is prioritized by GIFT, but not the other 2 hits on the left, even though these 2 hits had more significant marginal TWAS p-values.

For the panel about the region on Chr17, bottom left in **Fig. 3** (Fig. 4 in initial submission), GIFT did prioritize one significant gene RPRML.

5. - I have doubt that the APOC2 finding is true. Firstly, it seems the manuscript states that APOC2 encodes a protective apoE isoform, but APOC2 and APOE are different genes; this requires correction. More critically, the APOE locus is very complex and marked by strong linkage making it easy to obtain false positive variant or gene associations. Without a more careful and granular assessment of this finding and the

locus, the validity of this finding will remain a concern. Importantly also, the authors use Bellenguez et al. 2022 summary statistics, but those statistics are flawed at the APOE locus because P-values have been capped of at P<1e-300, which results in the APOE*4 variant, rs429358, and potential other variants being cut out of the summary statistics.

Response: We appreciate the reviewer's critical comment. Our initial wording incorrectly implied that *APOC2* encodes an apoE isoform, which is wrong. *APOC2* and *APOE* are distinct genes. We were meant to say that *APOC2* encodes a lipid-binding protein belonging to the apolipoprotein gene family, is a known associated gene of lipid traits, and located near the *APOE* locus.

In our analyses, *APOC2* was identified as a marginal TWAS-O signals in microglia (p-value = 1.41E-07) and oligodendrocytes (p-value = 1.15E-07). GIFT did not prioritize *APOC2* as an independent causal gene (see the GIFT results of Chr19, the region nearby *APOE*, in the bottom of **Fig. 3** and **Fig. S15**, and **Tables 2-4**). However, *APOC2* is validated in our PWAS-O findings. As shown in the PPI network **Fig4B**, *APOC2* is connected with *CLU* that is also a known AD risk gene. We have corrected this in the Results in lines292-298 on page 11.

The reviewer is also correct that the publicly available Bellenguez et al. 2022 summary statistics (that we used) did not include the most significant *APOE* E4 variants rs429358. In the *APOE* gene region (19:44905791-44909393), we considered 16 variants including two genome-wide significant variants rs769448 (19: 44906322; GWAS p-value = 7.81e-14) and rs769452 (19: 44907853; GWAS p-value = 7.494e-18).

However, the advantage of TWAS/PWAS is that all variants within +-1MB around the test gene are considered in the expression/protein prediction model. This would help still test the remain genetic effects in the *APOE* and its nearby genes. We have clarified this in the Discussion subsection "Limitations" on lines 448-452 on page 16, and the Methods subsection "GWAS summary data of AD dementia by Bellenguez et. al. 2022" in the bottom of page 24.

6. - The authors have compared their current results with their previous work (reference 20, https://doi.org/10.1186/s13195-025-01774-y). In the supplementary of this previous published paper, there were 113 significant bulk TWAS-O findings, which is inconsistent with the "Table S1. Bulk TWAS-O" of the current manuscript.

Response: We thank the reviewer for raising this point. We compared our cell-type-aware TWAS-O findings to the same bulk TWAS-O results in our previous work (**reference 20, T Hu et al. Alzheimer's Res Ther. 2025, https://doi.org/10.1186/s13195-025-01774-y**). The discrepancy of significant gene numbers is because different significance thresholds were considered in our previous work and this current paper.

In our previous work, significant bulk TWAS-O findings were defined using FDR q-value < 0.05, resulting in 113 significant genes. In this current manuscript, to maintain consistency with the cell-type-aware TWAS-O analyses and to adopt a more conservative genome-wide significance threshold, we define bulk TWAS-O significant using p-value < 2.5E-06. As shown

by the reduced total number of significant genes (shown in **Table S1**), this genome-wide significance threshold is more stringent than the FDR-based cutoff used previously.

We have updated the Results (lines 119-123 on page 5) and Supplementary Table S1 legend to clarify the employed statistical significance criterion.

7. - Adding captions for each Supplementary Tables will be helpful. In addition, for Tables S3–S8, please consider adding extra columns indicating whether each gene is validated by bulk TWAS-O or bulk PWAS-O. The additional columns could explicitly list the corresponding gene, chromosome, and bp position from the bulk analyses when overlap is detected.

Response: We thank the reviewer for this suggestion. In this revision, we have added detailed captions for all Supplementary Tables to improve clarity. For Tables S3-S8, we've added the additional columns for the overlapped significant genes by bulk TWAS-O or bulk PWAS-O. We also clarified in the captions that additional information about these overlapped bulk TWAS-O/PWAS-O risk genes could be found in Table S1-S2.

8. - Line 142-145, The description of bulk TWAS-O and bulk PWAS-O results is confusing. Please clarify explicitly that bulk TWAS-O comes from previously published work (ref. 20), whereas bulk PWAS-O was newly performed in this study. In addition, it would be helpful to indicate in the schematic (Figure 1) how these two bulk datasets were used for comparison.

Response: In the revised Results section, in lines 117-130 on pages 5-6, we have clarified that bulk TWAS-O results were taken from our previously published work (Ref. 20), which used the same bulk DLPFC RNA-seq data from the ROS/MAP cohort. We clarify that bulk PWAS-O was newly performed in this study using the same bulk DLPFC proteomic reference panel (**Fig. 1**). We further clarify that both bulk datasets were used exclusively for downstream comparison and validation and were not incorporated into the training or inference of cell-type-aware TWAS-O models. See cited revised texts as follows:

> *"We compared our cell-type-aware findings with two bulk-level analyses using reference omics data from the same ROS/MAP cohort and the same GWAS summary data of AD dementia[4]. Bulk TWAS-O results (**Table S1**) were obtained from our previously published work[20] using bulk DLPFC RNA-seq data (n=931). The total number of significant bulk TWAS-O risk genes shown in **Table S1** is less than the one reported in the previous publication (70 vs. 113), for using a more stringent genome-wide significance threshold p-values <2.5E-6 versus the FDR<0.05 use in the previous work. Bulk PWAS-O (**Table S2**) was newly performed in this study (n = 716).*
>
> *These bulk analyses were used exclusively for downstream comparison and validation and were not incorporated into model training for any cell-type-aware TWAS-O analyses. For each cell type, its cell-type-aware TWAS-O risk genes are considered as overlapping with bulk TWAS-O findings, when their test genetic regions overlapped. Independent cell-type-aware TWAS risk genes are considered as validated in PWAS-O findings, when their test genetic regions overlapped with any of the protein-coding genes with PWAS-O q-values<0.05."*

Minor Comments

9. - It's great to see QQ plots that indicate the inflation of TWAS/PWAS analyses, but it would be important to also report/add Genomic inflation factors so the readers can have a clearer sense of inflation.

Response: We thank the reviewer for this suggestion. We have now added the genomic inflation factors to the captions of the relevant Supplementary QQ plot figures, in **Fig. S5-S11**.

10. - Abbreviations should be checked for consistency. For example, "Alzheimer's disease" is written out multiple times after the abbreviation has already been introduced.

Response: We have carefully reviewed the manuscript and removed redundant instances where "Alzheimer's disease" was written out after the abbreviation "AD" had already been introduced. Abbreviation usage is now consistent.

Additional comments.

- Line 88, "cix-xQTL" should be "cis-xQTL".
Response: Fixed in this revision.

- In Table 1, the "2" in R^2 should be formatted as a superscript.
Response: Fixed in this revision.

- Fig .2, the y-axis label "-log10(q-value)" should be "-log10(p-value)"
Response: Fixed in this revision.

- Update your published paper (reference 20).
Response: Fixed in this revision.

- Line 264-265, "We identified 7 genes ……" is inconsistent with "ARL4A, RAB8B, HEXIM1, ZNF225, SNRPD2, and FBXO46" (6 genes).
Response: In the lines 257-259 on page 10 of the revised Results section, we have corrected this description as follows:

> *"Among risk genes validated in proteomics, 6 out of 7 excitatory neurons genes are specifically identified in excitatory neurons (ARL4A, RAB8B, HEXIM1, ZNF225, SNRPD2, and FBXO46)"*

- Many red dots were not labeled in Figure 2B-2F.
Response: We attempted to increase label coverage by adjusting the `max.overlaps` parameter in ggrepel. Due to severe label crowding in regions with many independent signals, some red dots remain unlabeled to preserve figure legibility. We now clarified this in the legend of Figure 2, see lines 920-924 on page 32, as follows:

> *"Independent significant genes are colored in red, including the ones detected by GIFT and the ones having no overlapped test regions with other significant TWAS-O risk genes. Most*

independent genes were labeled, with a few clustered with other independent genes were not labeled due to the limitation of the 'max.overlaps' parameter in the R plot function 'ggplot'."

- Line 544-550, please present the specific setting of generating PPI network in STRING.

Response: We have added the detailed STRING settings (i.e., version 12.0, minimum interaction score threshold 0.4) to the Methods section (lines 607-615) on page 22.

- In the abstract and main manuscript, please clarify that the 223 TWAS risk genes are unique genes derived from 350 significant gene–cell-type associations, to avoid ambiguity in the presentation.

Response: We now explicitly state in both the Abstract (line 31 on page 2) and the Results (line 78 on page 4) that 223 unique TWAS risk genes were identified from 350 significant associations, to avoid ambiguity between counts of genes and counts of associations.

- Line 145-147, it will be helpful to define the validation threshold of bulk TWAS-O.

Response: Added in this revision in lines 129-131 on page 6, the end of the first subsection in the Results part.

- Line 301, Line 848, and Line 857, it's not appropriate to use "physical", since many PPI links were not physical interaction.

Response: Fixed in this revision. We have avoided using "physical" to describe the PPI links in the Figure legends on page 32-33.

- Line 389, what is "in this previous study"?

Response: The ROS/MAP snRNAseq data we analyzed in this study were generated and first analyzed by Fujita, M. et. al., Nature Genetics 2024. Standard FUSION analyses were conducted for each of the cell type. In the revised Discussion, in lines 372-377 on page 14, we clarified the description by specifying the study and the data as follows:

"Compared to the previous cell-type-aware TWAS analyses by FUSION8 using the same ROS/MAP snRNA-seq and GWAS summary data19, we detected more significant TWAS risk genes using the more stringent genome-wide significance threshold of p-values<2.5E-06 (vs. FDR<0.05), 223 vs. 165 in the same 6 major brain cell types. Also, 40 out of our 91 independent cell-type-aware TWAS-O findings are shown overlapped with findings by this previous study by standard FUSION using the same ROS/MAP data."

- Figure 4 was not mentioned in the main manuscript and is identical to Figure S13.

Response: We changed the order of the Figure 4 in the initial submission to **Fig. 3** in this revision now. We have added brief descriptions about the example GIFT results in microglia in Fig. 3 and deleted the repeated one in the supplementary figures. See our revised texts in lines 149-153 on page 6 as follows:

"Example GIFT locus plots from microglia are shown in Fig. 3, where all genomic regions have at least one gene prioritized by GIFT. Because of joint modeling all nearby genes in the same model by GIFT, the prioritized genes are often not the genes with most significant marginal TWAS-O p-values in their test genomic regions."

- There are several inconsistencies need to be updated: Line 213 (20 & 16 flipped);

Line 179, “19” inconsistent with the counts in Table S3.
Response: These inconsistencies have been corrected as shown in line 155 and line 184 in the revised Results section, to be consistent with Supplementary Tables.

- In Figure S19, the labels bulk and proteome are potentially confusing. It would be clearer to explicitly indicate them as bulk TWAS-O and bulk PWAS-O, respectively, to avoid ambiguity.
Response: We have revised the labels to “bulk TWAS-O” and “bulk PWAS-O” in Figure S18 (previous Figure S19) to avoid ambiguity.

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Communications Biology initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

Response: We thank Reviewers #2 and #3 for their time and contribution to the peer-review process.

Reviewer #3 (Remarks to the Author):

This study by Liu et al. reports an omnibus methodology to identify cell type–specific risk genes. The authors utilized single-nuclei transcriptomic data to identify cell-type specific candidate genes and validated several associations through integration with bulk transcriptomics and proteomics data. They further employed PPI network data and pathway analysis to link these associations to functionally relevant biological pathways. Importantly, the study leverages one of the largest datasets currently available on Alzheimer’s disease (AD), reporting 11 novel associations. To strengthen the manuscript, I would encourage the authors to highlight not only the findings themselves but also their potential impact. While the methodology is improved, there are several limitations that should be addressed before publication.

Response: We appreciate this insightful suggestion. In the revised manuscript, we have added a new paragraph in the Discussion section that highlights the broader biological and translational significance of the 11 novel cell-type–aware genes identified in this study (Discussion, lines 425-438 on page 16) as follows:

> *“Beyond identifying risk genes in individual cell types, our results provide mechanistic insight into how genetic susceptibility to AD is distributed across the major neural and glial lineages. The 11 novel, GIFT fine-mapped independent genes discovered in this study highlight biological processes that may contribute to AD pathogenesis, including transcriptional regulation (MED18, KNOP1), chromatin remodeling and neuronal plasticity (KDM6B), oxidative stress responses (TP53INP1), cAMP signaling (PDE8B), estrogen biosynthesis (CYP19A1), and lipid-associated glial pathways (PLA2G2F, GRHL3). Importantly, these signals were not detectable in GWAS or bulk TWAS analyses, underscoring the added value of resolving genetically regulated expression*

at single-cell resolution. By integrating the results of TWAS-O, PWAS-O, fine-mapping (GIFT), our study delivers a prioritized set of cell-type-level candidate causal genes that extend current AD biology, and provide experimentally actionable entry points for mechanistic studies. Collectively, these findings refine the etiological map of AD risk by linking genetic effects to specific cellular contexts where pathogenic processes may be initiated or amplified."

1. A number of the reported cell type–specific risk genes appear to arise from non-specific expression. snRNAseq data is inherently prone to expression dropout and mis-annotation of poor quality cells. For example, one astrocyte-specific risk gene reported, CYP19A1, shows extremely low expression in astrocyte according to Human Protein Atlas data. Similarly, one of the genes validated through PWAS (C1QL1) is known to be myeloid cell-specific. Furthermore, GFAP, which is a canonical astrocyte marker, was reported here as a risk gene for oligodendrocytes. These examples suggest that the current analysis may have included spurious associations. I therefore recommend applying a minimum expression threshold or additional quality filters when deriving cell-specific associations from snRNA-seq data.

Response: We thank the reviewer for this thoughtful comment. The ROSMAP snRNA-seq data of DLPFC were profiled for 436 individual participants, resulting in a total of 1,546,794 cells. The sample size of 436 is much larger than previous cohort. Our analyses were conducted using pseudo bulk gene expression data quantified from the raw snRNA-seq data, which is expected to help with low expression in part of the single nucleus.

We carefully checked our data for these three highlighted genes: *CYP19A1*, *C1QL1*, and *GFAP*, using

(i) their snRNA-seq expression, mean of log2(CPM + 1) with pseudo bulk CPM, and the proportion samples with non-zero CPM.
(ii) cross-validated expression prediction performance (CV R^2) by TIGAR, PrediXcan, and FUSION tools.
(iii) the cis-eQTL weight patterns underlying the significant TWAS-O signals.

As shown in the following **Table 1**, the average expression level for *CYP19A1* is low, but our data did profile non-zero counts in >50% of the samples. The average expression levels for *C1QL1* and *GFAP* in the associated cell types are reasonable.

As shown in the following Table 2, the CV R2s for predicting the genetic components of these gene expression traits are low but still passed the 0.5% threshold we used to test the association. We now provide reference for using the CV R2 > 0.5% threshold to test the gene association in the beginning of the subsection of Results/Overview of the analytical procedure (lines 97-100 on page 5). Basically, as described in the Methods section, the gene-based association test in the Stage II of TWAS/PWAS is equivalent to a burden test with variant weights provided by their xQTL effect sizes. In the TIGAR-V2 paper, we showed that the TWAS/PWAS will increase power if the genetic effects of "xQTL" on phenotype are mediated through the expression/protein. However, if not, then the Stage II gene-based association test would just result in a non-significant p-values as a test under the NULL hypothesis. This is why the TWAS/PWAS tests are still statistically valid even if the CV R2 is relatively low.

As shown in the following **Fig. 1**, the highlighted test "xQTL" with marginal GWAS p-values < 1E-5 are believed to be the genetic variants contributing to the significant gene-based TWAS/PWAS association tests.

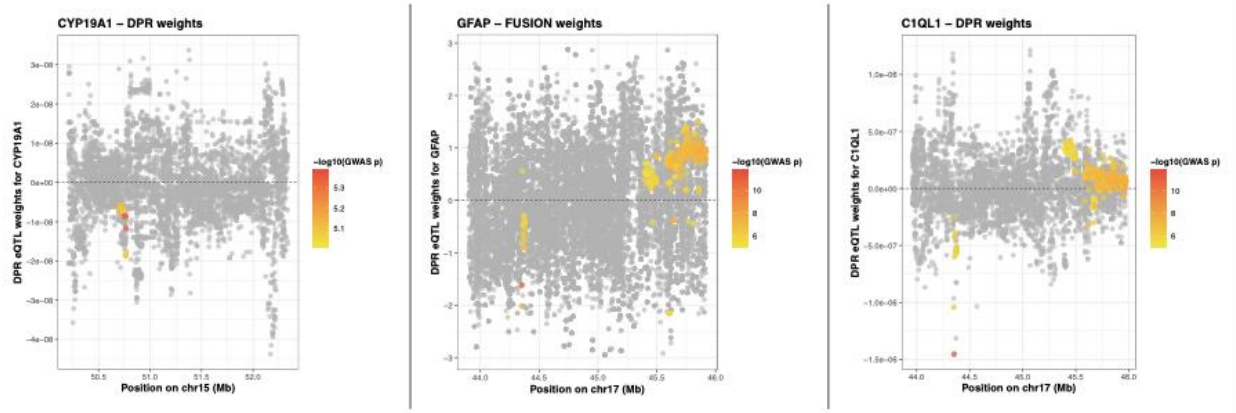
Table 1. Expression profiles of example genes

Gene	Cell type	Mean log2(CPM+1) of pseudo bulk expression data	Proportion of Samples with Non-zero CPM
CYP19A1	Astrocytes	0.015147116	61.47%
CYP19A1	Excitatory Neurons	0.014255199	86.24%
CYP19A1	Microglia	0.056722573	53.40%
C1QL1	Astrocytes	0.458197053	97.71%
C1QL1	Microglia	1.689502942	99.31%
GFAP	Astrocytes	3.552743455	98.85%
GFAP	Oligodendrocytes	5.294999070	100.00%

Table 2. Cross-validated prediction performance of example genes across three model

Gene	Cell type	Method	CV R^2	TWAS p-value
CYP19A1	Astrocytes	TIGAR/DPR	0.00506637	6.83E-07
CYP19A1	Astrocytes	PrediXcan/EN	-	-
CYP19A1	Astrocytes	FUSION	-	-
C1QL1	Astrocytes	TIGAR/DPR	0.00744699	4.72E-07
C1QL1	Astrocytes	PrediXcan/EN	0.00591065	-
C1QL1	Astrocytes	FUSION	-	-
GFAP	Oligodendrocytes	TIGAR/DPR	-	-
GFAP	Oligodendrocytes	PrediXcan/EN	-	-
GFAP	Oligodendrocytes	FUSION	0.00735144	1.37E-08

Fig 1. eQTL weights plots of example genes for method that gives significant p-values.



2. In a recently published study (Haglund et al., https://www.nature.com/articles/s41588-024-02050-9) , a different cohort of snRNA-seq data was used to map cell type–specific eQTLs. Comparing the current findings against this independent dataset would strengthen the confidence in the identified targets.

Response: We thank the reviewer for this constructive suggestion. We compared our cell-type-aware TWAS-O findings against the cell-type-specific cis-eQTLs reported by Haglund et al. (Nat Genet 2024. PMID: 39794547), which were generated from an independent snRNA-seq cohort of 391 individuals of European ancestry. We applied the genome-wide significance threshold of P < 2.5E-06 to filter their results and examined the proportion of overlapped eGenes and our TWAS-O risk genes (**Table 3** as follows). We added the following descriptions of this comparison in the first paragraph of the revised Discussion section (lines 363-368 on page 14).

> *"Comparing to the recently published cell-type-aware cis-eQTL mapping[68] derived from an independent snRNA-seq cohort of European ancestry, we found that 6%-17% of our TWAS-O risk genes across all six brain cell types were identified as eGenes containing at least one significant cis-eQTL. This is expected as all cis-SNPs with non-zero effect sizes on the target gene expression, often not statistically significant cis-eQTL, are tested by TWAS-O."*

Table 3. Summary of overlap results

Cell Type	The number (proportion) of TWAS-O risk genes that are also eGenes
Astrocytes	11 of 65 (16.9%)
Excitatory Neurons	11 of 77 (14.3%)
Inhibitory Neurons	6 of 49 (12.2%)
Microglia	11 of 64 (16.9%)
Oligodendrocytes	11 of 50 (22.0%)
Oligodendrocyte Precursor Cells	4 of 45 (8.9%)

3. The pragmatic choice of CV R2>0.5% as a threshold is understandable given sample size limitations. However, such a low threshold also risks including weakly predicted genes. I recommend testing higher cutoffs (e.g. 1% or 5%) and reporting how sensitive the results are to this choice, as this would provide readers with a clearer picture of the robustness of the findings.

Response: We thank the reviewer for the comment. The use of CV $R^2 > 0.5\%$ follows the recommendation from our previous TIGAR-V2 paper (RL Parrish et al. HGG Adv. 2021, PMID: 35047855), which demonstrated that a liberal threshold allows inclusion of genes with modest heritability in snRNA-seq datasets without inflating type I errors. We also explained this in our above response to the comment 1 of Reviewer 3, cited as follows:

> "We now provide reference for using the CV R2 > 0.5% threshold to test the gene association in the beginning of the subsection of Results/Overview of the analytical procedure (lines 97-100 on page 5). Basically, as described in the Methods section, the gene-based association test in the Stage II of TWAS/PWAS is equivalent to a burden test with variant weights provided by their xQTL effect sizes. In the TIGAR-V2 paper, we showed that the TWAS/PWAS will increase power if the genetic effects of "xQTL" on phenotype are mediated through the expression/protein. However, if not, then the Stage II gene-based association test would just result in a non-significant p-values as a test under the NULL hypothesis. This is why the TWAS/PWAS tests are still statistically valid even if the CV R2 is relatively low."

To evaluate robustness, we performed a sensitivity analysis using more stringent cutoffs (CV $R^2 > 1\%$, the threshold commonly used in the field). As expected, the number of valid imputation models decreased substantially across all cell types due to the sparse nature of snRNA-seq expression. However, the majority of fine-mapped independent TWAS-O risk genes (78-92% across cell types) remained significant under CV $R^2 > 1\%$, and all 11 novel fine-mapped genes remained detectable under both alternative thresholds. These results indicate that while a more stringent CV R^2 threshold reduces the number of testable genes, the key biological conclusions of this study are stable across thresholds. These results demonstrate that our primary biological conclusions are not sensitive to different prediction-performance thresholds.

Table 4. Sensitivity of TWAS-O independent risk genes to prediction-accuracy threshold

Cell Type	**The number of tested genes with valid models under CVR^2>1% (% of current significant associations were still identified)**	**Independent TWAS-O genes retained**
Astrocytes	DPR: 9078 of 13482 (67.3%) EN: 2670 of 4536 (58.9%) FUSION: 2799 of 5478 (51.1%)	10 of 15 (66.7%)
Excitatory Neurons	DPR: 9550 of 13649 (70.0%) EN: 3274 of 4929 (66.4%) FUSION: 2799 of 5478 (51.1%)	14 of 22 (63.6%)

Inhibitory Neurons	DPR: 8888 of 13389 (66.4%) EN: 2449 of 4239 (57.8%) FUSION: 2562 of 5176 (49.5%)	12 of 15 (80.0%)
Microglia	DPR: 8826 of 13458 (65.6%) EN: 2248 of 4126 (54.5%) FUSION: 2370 of 5017 (47.2%)	21 of 23 (91.3%)
Oligodendrocytes	DPR: 8990 of 13462 (66.8%) EN: 2415 of 4155 (58.1%) FUSION: 2555 of 5138 (49.7%)	12 of 17 (70.6%)
Oligodendrocyte Precursor Cells	DPR: 8534 of 13282 (64.3%) EN: 2091 of 3898 (52.1%) FUSION: 2352 of 4710 (49.9%)	10 of 14 (71.4%)

4. As per my understanding, none of the xWAS method directly tests for colocalization. Therefore, I also recommend performing colocalization analyses to help distinguish true causal relationships from signals driven by linkage disequilibrium. This would substantially increase the mechanistic interpretability of the results.

Response: We thank the reviewer for raising the important point regarding colocalization. First, we want to explain why we do not consider COLOC as an alternative computational validation tool for our TWAS-O findings. This is because the goal of the COLOC tool (Wallace et al. 2021. PLOS Genetics; PMC: 8504726) is to consider if the "H4: both expression and disease are associated with the same single causal variant" would account for most probabilities (the threshold of >0.75 is often used). However, as shown in the detailed Methods description, TWAS is to test if the combined weighted genetic effect of all cis-SNPs with non-zero xQTL effect sizes (in the expression/protein prediction model) is associated with the disease. It is the xQTL effect size weighted combination of the marginal GWAS Z-scores of all cis-SNPs that leads to powerful TWAS/PWAS analyses.

Given summary-level eQTL and GWAS data of a test genomic region, the recently released COLOC tool will first fine-map significant eQTL and GWAS signals by using the SuSiE fine-mapping method, and then calculate the probabilities of the "H4: both expression and disease are associated with the same single causal variant" for each pair of the fine-mapped few significant eQTL and GWAS signals. See the following description from the COLOC paper:

> "*Explicitly, if L1 and L2 signals are detected (have a credible set returned) for traits 1 (e.g., expression) and 2 (disease) respectively, then the colocalization algorithm is run L1 × L2 times. Thus, the user is presented with two lists of signals for each trait, and the L1 × L2 matrix of pairwise posterior probabilities of H4 may be examined to infer which pair of tags, if any, represent the same signal.*"

Alternatively, we applied the recently proposed Probability Mendelian Randomization method, PMR-Egger, which is tailored for gene-level causal inference. We provided descriptions of the PMR-Egger method in the Methods/Causal inference using PMR-Egger subsection, lines 591-605 on page 22. We showed that the PMR-Egger assumes the same model framework as the TWAS/PWAS framework, testing for the causal genetic effects mediated through

expression/protein. The differences are that PMR-Egger modeles the separate two stages of TWAS/PWAS in one statistical framework, and specifically models the horizontal pleiotropy effects between expression/protein and disease. The disadvantage of PMR-Egger method is its computation burden for testing each gene. Thus, we applied PMR-Egger to all of our significant cell-type-aware TWAS-O risk genes in six cell types (see results in the following **Table 5**).

In this revision, we included the significant PMR-Egger FDRs in **Tables S3-S8**, and added the following paragraph in the Results section, lines 189-198 on page 8, to describe the alternative computational validation by PMR-Egger:

> *"To further evaluate whether the significant TWAS-O associations reflect causal genetic effects mediated through gene expression, we applied the probabilistic Mendelian randomization method PMR-Egger36 to all cell-type-aware significant genes. Significant PMR-Egger FDRs are presented in Tables S3-S8. We found that 34% of the of the total 350 significant TWAS-O associations in six cell types show significant PMR-Egger FDR<0.05. A substantial proportion of TWAS-O independent signals showed significant mediated causal genetic effects by PMR-Egger: 9 of 15 in astrocytes, 16 of 23 in microglia, 15 of 22 in excitatory neurons, 10 of 15 in inhibitory neurons, 10 of 17 in oligodendrocytes, and 8 of 14 in OPCs. These results provide a computational validation of our identified cell-type-aware risk genes of AD dementia."*

Table 5. PMR-Egger results

Cell Type	**The number of PMR-Egger significant genes out of the total TWAS-O significant genes**	**The number of PMR-Egger significant genes overlapping with independent TWAS-O genes**
Astrocytes	21 of 65 (32.3%)	9 of 15 (60.0%)
Excitatory Neurons	27 of 77 (35.1%)	15 of 22 (68.2%)
Inhibitory Neurons	19 of 49 (38.8%)	10 of 15 (66.7%)
Microglia	25 of 64 (39.1%)	16 of 23 (69.6%)
Oligodendrocytes	15 of 50 (30.0%)	10 of 17 (58.8%)
Oligodendrocyte Precursor Cells	12 of 45 (26.7%)	8 of 14 (57.1%)

Responses to Reviewers' Comments

Reviewer #1 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the comments on behalf of both of us.

Reviewer #2 (Remarks to the Author):

I thank the authors for their detailed and thoughtful responses. They have made substantial efforts to address the reviewers' comments, and most concerns have been adequately resolved. Only one major issue remains that should be addressed prior to publication, while two additional points would benefit from minor clarification to further improve the manuscript.

Major (strongly recommended to be addressed before publication)

Without explicitly conditioning on APOE*4 and APOE*2 allele dosage, reliable gene-level inference at the APOE locus is not possible due to extreme linkage disequilibrium and dominant APOE-driven effects.

Consequently, the conclusions regarding APOC2 are not sufficiently supported.
The authors should therefore either adjust for apoe4 and apoe2 dosage for analyses within the APOE locus or **completely remove the locus and related discussion from the manuscript. (Only in discussion).**

Response: We thank the reviewer for this important clarification. We agree that gene-level inference within the APOE locus is challenging due to exceptionally strong linkage disequilibrium and dominant haplotype-driven effects from *APOE*4* and *APOE*2*, and that the Bellenguez 2022 GWAS summary data excluded the *APOE*4/2* locus. To avoid over-interpretation, in the revised manuscript we removed the interpretation about the *APOC2* results from the Results section and downstream interpretation (subsection "**PPI network and enrichment analyses results**", lines 290-293 on page 11). We additionally added an explicit Discussion paragraph (subsection "**Limitations**", lines 453-462 on pages 16-17) explaining why the *APOE* locus and nearby TWAS/PWAS risk genes were excluded in our results interpretations, cited as follows:

> *"Second, the publicly available GWAS summary statistics used in this study excluded variants with significant p-values<1E-300 in the APOE locus, including the well-known risk variant APOE E2 rs7412 and E4 rs429358. Because TWAS/PWAS framework would test all available variants within ±1MB region around the test gene, the remaining genetic effects were still tested for gene APOE and its nearby genes. However, the APOE region is characterized by extremely strong linkage disequilibrium with the dominant APOE E2/E4 GWAS loci, making the nearby TWAS/PWAS risk genes difficult to interpret without explicitly conditioning on APOE E2/E4. To avoid over-interpretation, we therefore excluded the APOE locus and nearby TWAS/PWAS risk genes from downstream interpretations."*

Minor (partially addressed; suggested minor revisions)

Line 539–540. The revised wording remains potentially misleading. The statement that "DPR yields substantially more valid expression prediction models, particularly for genes with relatively low expression heritability" could be interpreted as implying improved predictive reliability. In practice, DPR primarily increases model permissiveness and the number of fitted models, rather than predictive accuracy per se. A slight rephrasing to clarify this distinction would further improve accuracy and clarity.

Response: We thank the reviewer for this clarification. We have revised the sentences to explicitly clarify that DPR yields a larger number of expression prediction models with CV R2>0.5%, a criterion used to select genes/proteins for the Stage II gene-based association tests (see subsection "**Motivation for including TIGAR/DPR in xWAS-O**", lines 549-552 on page 20).

Line 372–377. This comparison has been partially clarified, but it may still be misleading. The increased number of significant findings is likely driven by the substantially larger set of genes tested using TIGAR/DPR-based models, rather than by the use of a more stringent significance threshold itself. Explicitly stating this would help avoid overinterpretation and further strengthen the presentation.

Response: We thank the reviewer for this suggestion. We have revised the text to explicitly state that the difference is likely due to the larger set of tested genes by TIGAR/DPR method and the power gained by aggregating TWAS p-values based on three complementary statistical method (by ACAT method), rather than implying that a more stringent threshold yields more discoveries. (see subsection "**Alternative studies validated cell-type-aware TWAS-O findings**", lines 378-381 on page 14).

Reviewer #3 (Remarks to the Author):

Thanks to authors for addressing the concerns appropriately. The PPI network and pathway figures (figure 4b,c and 5 are distorted and not aesthetically pleasant. I recommend to properly format the figures suitable for publication.

Response: We thank the reviewer for this suggestion. We have reformatted both PPI network and pathway enrichment figures to improve readability and publication quality (**Fig. 4b, c** and **Fig.5**).

We simplified the network visualization by encoding the STRING evidence support as edge thickness (STRING confidence score), rather than displaying multiple evidence categories simultaneously, which improves clarity while preserving the interpretation that edges represent functional protein-protein associations. Accordingly, we updated the Methods and figure legends to reflect this notation (subsections "Protein-protein interaction network analyses by STRING", lines 619-627 on pages 22-23 and "Figure legends", lines 952-954, 959-961 on pages 34).