Commentary | Open Access

Journal of Translational Genetics and Genomics

Wei *et al. J Transl Genet Genom.* 2026;10:645-51 **DOI:**10.20517/jtgg.2026.56

Multiomics integration and causal inference in aging research—can the molecular gap be bridged?

Liuxiang Wei[1], Xiaolin Ni[2]

Citation: Wei L, Ni X. Multiomics integration and causal inference in aging research—can the molecular gap be bridged?. *J Transl Genet Genom.* 2026;10:645-51. https://dx.doi.org/10.20517/jtgg.2026.56

Received: 8 May 2026 **First Decision:** 22 Sep 2026 **Revised:** 13 Oct 2026 **Accepted:** 22 Oct 2026 **Published:** 29 Oct 2026

Academic Editor: Wenzhao Wang **Copy Editor:** Fangling Lan **Production Editor:** Fangling Lan

Check for updates

OVERVIEW

Aging constitutes a multifaceted biological process influenced by the interplay of numerous factors. It encompasses genetic variation, transcriptional regulation, epigenetic modifications, protein synthesis, signaling pathways, and the functional integrity of tissues and organs, as well as environmental risk factors[1,2]. Contemporary research on aging has increasingly transitioned from focusing on isolated components to employing multi-omics approaches that span molecular, cellular, tissue, organ, and even individual phenotypic levels[3]. The primary objective of these investigations is to elucidate the molecular mechanisms underlying aging and age-associated diseases, with the ultimate goal of mitigating or preventing the onset of aging and related pathologies, thereby promoting health and longevity[4-6]. Among the available molecular layers, expression quantitative trait locis (eQTLs) capture genetic effects on transcript abundance; splicing uantitative trait locis (sQTLs) capture effects on RNA splicing; alternative polyadenylation uantitative trait locis (apaQTLs) capture effects on 3' UTR usage; protein uantitative trait locis (pQTLs) capture effects on circulating protein levels; and metabolite uantitative trait locis (mQTLs) capture effects on metabolite abundance[7]. Each layer provides complementary regulatory information, but none alone can capture the full architecture of aging, given tissue specificity, platform differences, sample size limits, and ancestry bias. Integrating multiple quantitative trait loci (QTL) layers with aging-related genome-wide association study (GWAS) data can therefore increase statistical power, help distinguish shared causal variants from linkage disequilibrium (LD)-driven associations, and link genetic variation to molecular phenotypes and aging outcomes, although such integration remains hypothesis-generating and requires functional validation. Gai *et al.* have conducted one of the most extensive multi-omics integration studies to date, examining five distinct layers of molecular QTLs in relation to epigenetic aging acceleration and human longevity[8]. Employing analytical frameworks such as Generalized Summary-based Mendelian Randomization (GSMR), Summary-based Mendelian Randomization (SMR),

[1]Jiangbin Hospital of Guangxi Zhuang Autonomous Region, Nanning 530021, Guangxi, China.
[2]Institute of Basic Medical Sciences & School of Basic Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China.

Correspondence to: Prof. Xiaolin Ni, Institute of Basic Medical Sciences & School of Basic Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100005, China. E-mail: xiaolin_ni7@126.com

Heterogeneity In Dependent Instruments (HEIDI), and Omics Pleiotropic Association (OPERA), their investigation spanned multiple biological levels, including gene function, transcriptional expression, epigenetic regulation, protein translation, signaling pathways, tissue specificity, and disease associations. Their integrative approach aimed to elucidate the underlying molecular mechanisms of aging and identify potential pharmacological targets for therapeutic intervention. Compared with recent metabolomic and single-cell aging studies[4-6], Gai *et al.* provide a unique multi-QTL integration focused on epigenetic aging and longevity, combined with drug-target annotation[8]. However, their approach is limited by the availability of bulk-tissue QTL data and the lack of single-cell and spatial resolution.

Integration of under-studied molecular layers

Building on this multi-layer framework, previous research on aging-related QTL has predominantly concentrated on eQTLs[7]. In the present study, the integration of sQTLs, apaQTLs, pQTLs, and mQTLs uncovers previously unrecognized biological insights. In total, Gai *et al.* identified 30 genes, 11 splicing events, 5 proteins, 3 APA events, and 39 metabolites associated with aging-related traits, including the metabolite oleoyl-linoleoyl-glycerol, which was negatively associated with longevity and showed unfavorable relationships with high-density lipoprotein and low-density lipoprotein[8]. The OPERA analysis further identified the microtubule-associated protein tau locus, a region well known for its association with neurodegenerative diseases, as a colocalization locus involving three types of molecular phenotypes (posterior probability of association > 0.7), suggesting that its regulatory landscape may also influence longevity and epigenetic aging. Cross-tissue analysis reported by Gai *et al.*[8] revealed that CASP8 was positively associated with longevity in the spleen, lung, and heart, and secretagogin (SCGN) showed a significant positive SMR association with Intrinsic Epigenetic Age Acceleration in adipose and thyroid tissues, with age-associated downregulation of SCGN expression. SORT1 was positively associated with longevity specifically in the liver but not in whole blood, highlighting the importance of tissue context. Of particular significance, the gene FBXL5 demonstrated a robust association with longevity exclusively through apaQTLs, with no corresponding association observed via eQTLs. This finding suggests a potential role for alternative polyadenylation as a regulatory mechanism in lifespan determination that has hitherto been insufficiently acknowledged, although functional validation is required. The absence of a corresponding eQTL signal does not by itself prove APA-mediated causality, and experimental approaches such as 3' RACE, APA reporter assays, clustered regularly interspaced short palindromic repeats (CRISPR) editing of the polyadenylation signal, and RNA stability measurements are needed to validate this hypothesis. Consistent with this, APA-mediated 3' UTR lengthening has been observed in cellular senescence, and APA is increasingly recognized as a key modulator of gene expression in age-related pathologies[9]. It is important to note that no single QTL layer can capture the full regulatory architecture of aging. For example, eQTLs are often tissue-specific and may not reflect protein abundance or post-transcriptional regulation[10]; sQTLs and apaQTLs are highly context-dependent and remain under-studied in aging[7]; pQTLs and mQTLs are influenced by sample source, platform, and ancestry[11]. In addition, these QTL datasets differ substantially in sample size, age distribution, sex composition, genetic ancestry, tissue or cell type, and experimental platform[8,12]. For example, eQTLGen is based on whole blood (n = 31,684), BrainMeta on brain (n = 2,865), GTEx apaQTL on whole blood (n = 670), deCODE pQTL on plasma (n = 35,559), and Canadian longitudinal study on aging mQTL on metabolomics (n = 8,299). Most datasets are predominantly of European ancestry and are not uniformly designed for aging-specific analyses[13]. Such heterogeneity can affect LD structure, instrument strength, winner's curse, sample overlap, tissue-specific causal inference, and external validity, thereby complicating cross-layer integration and the portability of causal inference[14,15]. Nevertheless, these cross-layer signals provide valuable candidate directions for future functional and epidemiological studies.

Causal inference across molecular QTL layers: a tiered framework and its caveats

In molecular epidemiological investigations of longevity and epigenetic aging, a primary challenge lies in differentiating putative causal relationships from confounding factors, reverse causation, or spurious signals

arising from LD imbalance[16,17]. In this context, Gai *et al.* have made notable methodological contributions by systematically integrating multi-level molecular QTL data with aging-related GWAS datasets, and by implementing a tiered statistical causal inference framework[8].

Initially, the GSMR approach was applied, incorporating LD clustering and the removal of HEIDI outliers to minimize the confounding effects of horizontal pleiotropy[18,19]. Subsequently, for each significant association identified, up to nine alternative Mendelian randomization (MR) methods were employed for cross-validation. This multi-method consensus approach substantially mitigates the risk of false positives that may arise from reliance on a single MR model assumption[19,20]. Of particular note is the stringent criterion applied in the SMR analysis, wherein genes and proteins were retained only if they satisfied a HEIDI test p-value greater than 0.01, thereby effectively distinguishing genuine pleiotropic effects from chance associations driven solely by LD[18].

This tiered analytical framework—progressing from genetic correlation assessment, through primary GSMR analysis, to sensitivity verification via multiple MR methods, and culminating in HEIDI filtering—establishes a rigorous analytical paradigm for aging research. It enhances the causal interpretability and reliability of subsequently identified candidate targets, including 30 genes, 11 splicing events, and 39 metabolites, surpassing the inferential power of simple correlation or co-localization analyses. However, these methods rest on several strong assumptions. GSMR and SMR require genetic variants to be valid instruments—associated with the molecular phenotype, independent of confounders, and affecting the outcome only through the phenotype. Horizontal pleiotropy, LD, winner's curse, and sample overlap can violate these assumptions and bias effect estimates[14,15,21]; for instance, simulations have shown that GSMR type I error rates are inflated when sample overlap increases, and GSMR may ignore weak pleiotropic effects, potentially leading to inflated type I errors[22]. HEIDI reduces LD-driven false positives but can be conservative and may not fully exclude pleiotropy; the test can output NA when its assumptions are violated, and its power to distinguish causality from pleiotropy is limited when causal variants are in close to perfect LD[23]. OPERA's posterior probability of association depends on model priors and should be viewed as evidence for shared causal variants rather than proof of causality; the associations identified can occur even if the causal variants for the molecular phenotypes are distinct from, but in LD with, the causal variants for the trait, and OPERA only provides gene-level tests without pinpointing specific variants[23]. Moreover, tissue heterogeneity and limited ancestry diversity may reduce the portability of these inferences. It should also be noted that the layer-specific findings reported above, including 30 genes, 11 splicing events, and 39 metabolites, were derived from individual SMR analyses rather than being jointly supported by multiple QTL layers. Multi-layer-supported signals were identified by the OPERA analysis, which detected 44 longevity-associated loci with joint evidence from at least two molecular QTL layers[8]. In addition, this analytical hierarchy does not imply a sequential biological mediation pathway across molecular layers. Formal mediation analysis, longitudinal multi-omics data and experimental perturbation are required to establish sequential mediation. Nonetheless, it is important to acknowledge that this methodology cannot eliminate the influence of unknown horizontal pleiotropy or biases introduced by sample overlap[22,24]. Despite these limitations, the approach represents one of the most robust statistical causal inference frameworks currently achievable given existing data constraints.

Mechanistic gaps will require further investigation

Through the application of SMR and OPERA analyses, genes including CASP8, PSRC1, and SORT1 were identified[8]. Notably, CASP8, traditionally recognized for its role in apoptosis, exhibits a positive correlation with longevity, suggesting potential non-apoptotic functions, possibly related to immune homeostasis or cellular quality control[25]. The liver-specific association of PSRC1 and SORT1 further underscores the importance of tissue-specific regulation in longevity. While these findings are significant, functional

validation remains essential. As Gai *et al*. have noted[8], integrative approaches based on xQTL data cannot substitute for direct experimental interventions. For instance, the identification of an apaQTL signal affecting FBXL5, absent in eQTL datasets, prompts the question of how alternative polyadenylation of this gene may influence lifespan. Addressing such mechanistic gaps will require further investigation using model organisms or cellular systems. For instance, CRISPR-Cas9 or prime editing of the 3' UTR polyadenylation signal in hepatic organoids or hepatocyte cell lines (e.g., HepG2 or Huh7), followed by 3' RACE, APA reporter assays, and RNA stability measurements, as well as mouse models with altered FBXL5 APA isoform usage. To resolve tissue-specific signals, single-cell RNA sequencing and spatial transcriptomics in liver, brain, blood, adipose, and thyroid tissues could identify the genes of cell types driving tissue-specific QTL effects[26]. In addition, the sample size and ancestry limitations of the QTL datasets used here—most of which are predominantly European and based on whole blood, plasma, or bulk brain tissue—should be addressed by generating larger, cross-ancestry, and tissue-specific QTL datasets.

Enhancement of the association map linking pharmaceutical compounds and their biological targets

The mapping of associations between drugs and their molecular targets constitutes a critical component in longevity research, bridging the gap from molecular discovery to therapeutic intervention[27]. Gai *et al*. employed the DrugBank database to link genes and proteins identified through SMR analyses to established or experimental pharmacological agents, thereby offering direct insights for drug repurposing efforts[8]. For example, NFKB1, a central regulator of inflammatory processes, has been reported as a potential mediator of aspirin's anti-inflammatory effects in atherosclerosis, although this evidence is context-dependent and requires further validation[28]; similarly, the interaction between feline sarcoma oncogene and fostamatinib suggests potential therapeutic avenues for modulating immune-related aging pathways[29]. The principal advantage of this supplementary approach lies in its capacity not only to catalog molecular phenotypes but also to actively associate these with existing drug resources, thereby expediting the translation from genetic findings to clinical evaluation. This is particularly pertinent for targets such as CASP8 and PSRC1, which exhibit significant associations with longevity[8]; the presence of drugs targeting their homologous proteins or related pathways could prioritize these candidates for experimental validation or dosage optimization. Nonetheless, the current drug-target association map predominantly reflects correlative relationships and lacks critical annotations regarding tissue specificity, dosage effects, and adverse reactions[30-32]. Moreover, most associations derive from literature reports or structural similarity assessments rather than direct in vivo intervention evidence[33,34]. Additionally, the small-molecule drugs cataloged in DrugBank primarily target enzymes or receptors, with minimal representation of splicing events or mQTLs[35]. Consequently, while this supplementary strategy offers valuable hypotheses, it should be considered a tool for generating assumptions rather than a definitive therapeutic framework[36,37]. Future integration of drug-target associations with tissue-specific expression profiles, age-related expression dynamics, and multi-omics co-localization data, together with in vivo intervention evidence, holds promise for more effectively guiding precision intervention trials in longevity research.

CONCLUSION

Gai *et al*. synthesized data from five molecular QTL studies alongside aging-related GWAS to develop a comprehensive molecular map of longevity and epigenetic aging[8]. This work is distinguished by three principal contributions: first, the establishment of a multi-tiered statistical causal inference framework that explores potential causal links from genetic variants to molecular phenotypes and subsequently to aging-related outcomes, thereby substantially improving the directionality and robustness of association analyses under MR assumptions; second, the identification of the potential involvement of non-coding regulatory mechanisms, such as alternative polyadenylation, in the aging process, thus expanding the scope of aging research; and third, the generation of a candidate list of potential targets for drug repurposing through drug-target mapping, effectively bridging the gap between fundamental biological discoveries and clinical

application. Future investigations should prioritize several key directions: increasing the sample sizes of epigenetic and metabolomic GWAS and developing cross-population QTL datasets; incorporating single-cell and spatial transcriptomic methodologies to elucidate cell type-specific regulatory mechanisms implicated in aging; and performing both in vitro and in vivo functional assays to test the putative causal roles of candidate targets identified herein, including but not limited to CASP8, PSRC1, and FBXL5. A profound integration of computational and experimental biology is essential to translate "longevity-associated molecules" into viable "anti-aging therapeutic targets".

DECLARATIONS

Authors' contributions

Contributed to the conception and writing of the manuscript: Wei L, Ni X

Availability of data and materials

Not applicable.

AI and AI-assisted tools statement

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

Both authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2026.

REFERENCES

1. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. Hallmarks of aging: an expanding universe. *Cell.* 2023;186:243-78. DOI PubMed
2. Moqri M, Herzog C, Poganik JR, et al. Biomarkers of aging for the identification and evaluation of longevity interventions. *Cell.* 2023;186:3758-75. DOI PubMed PMC
3. Chen C, Wang J, Pan D, et al. Applications of multi-omics analysis in human diseases. *MedComm.* 2023;4:e315. DOI PubMed PMC
4. Zhang S, Wang Z, Wang Y, et al. A metabolomic profile of biological aging in 250,341 individuals from the UK biobank. *Nat Commun.* 2024;15:8081. DOI PubMed PMC
5. Li G, Zhu X, Xiao F, et al. A multi-omics molecular landscape of 30 tissues in aging female rhesus macaques. *Nat Methods.* 2025;22:2658-69. DOI
6. Li Y, Wang Q, Xuan Y, et al. Investigation of human aging at the single-cell level. *Ageing Res Rev.* 2024;101:102530. DOI
7. Wen J, Nodzak C, Shi X. QTL analysis beyond eQTLs. In: Shi XM, editor. eQTL analysis. New York, NY: Springer US; 2019. pp. 201-10. DOI
8. Gai K, Li W, Li J, et al. Genetic and molecular factors underlying human longevity and epigenetic aging. *NPJ Aging.* 2026;12:84. DOI PubMed PMC
9. Chen M, Zhu X, Hang L, et al. Alternative polyadenylation in aging and aging-related diseases. *Aging Dis.* 2026. DOI
10. Zhang J, Zhao H. eQTL studies: from bulk tissues to single cells. *J Genet Genomics.* 2023;50:925-33. DOI PubMed PMC
11. Yang C, Gorijala P, Timsina J, et al. European and African ancestry-specific plasma protein-QTL and metabolite-QTL analyses identify ancestry-specific T2D effector proteins and metabolites. *Nat Commun.* 2025;16:7412. DOI PubMed PMC

12. Aguet F, Anand S, Ardlie KG, et al. The GTEx consortium atlas of genetic regulatory effects across human tissues. *Science.* 2020;369:1318-30. DOI
13. Long E, García-Closas M, Chanock SJ, Camargo MC, Banovich NE, Choi J. The case for increasing diversity in tissue-based functional genomics datasets to understand human disease susceptibility. *Nat Commun.* 2022;13:2907. DOI PubMed PMC
14. Forde A, Hemani G, Ferguson J. Simulated sample splitting approach to address biases due to instrument selection and participant overlap in two-sample Mendelian Randomization studies. *PLoS Genet.* 2026;22:e1011949. DOI PubMed PMC
15. Wang K. Interval estimate of causal effect in summary data based Mendelian randomization in the presence of winner's curse. *Genet Epidemiol.* 2024;48:74-84. DOI PubMed
16. Liu Z, Shen J, Barfield R, Schwartz J, Baccarelli AA, Lin X. Large-scale hypothesis testing for causal mediation effects with applications in genome-wide epigenetic studies. *J Am Stat Assoc.* 2021;117:67-81. DOI PubMed PMC
17. Yashin AI, Zhbannikov I, Arbeeva L, et al. Pure and confounded effects of causal SNPs on longevity: insights for proper interpretation of research findings in GWAS of populations with different genetic structures. *Front Genet.* 2016;7:188. DOI PubMed PMC
18. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet.* 2016;48:481-7. DOI
19. Zhu Z, Zheng Z, Zhang F, et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat Commun.* 2018;9:224. DOI PubMed PMC
20. Burgess S, Davey Smith G, Davies NM, et al. Guidelines for performing Mendelian randomization investigations: update for summer 2023. *Wellcome Open Res.* 2023;4:186. DOI PubMed PMC
21. Jiang T, Gill D, Butterworth AS, Burgess S. An empirical investigation into the impact of winner's curse on estimates from Mendelian randomization. *Int J Epidemiol.* 2023;52:1209-19. DOI PubMed PMC
22. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol.* 2016;40:597-608. DOI PubMed PMC
23. Wu Y, Qi T, Wray NR, Visscher PM, Zeng J, Yang J. Joint analysis of GWAS and multi-omics QTL summary statistics reveals a large fraction of GWAS signals shared with molecular phenotypes. *Cell Genomics.* 2023;3:100344. DOI PubMed PMC
24. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44:512-25. DOI PubMed PMC
25. Graham AC, Bellou E, Harwood JC, et al. Human longevity and Alzheimer's disease variants act via microglia and oligodendrocyte gene networks. *Brain.* 2025;148:969-84. DOI PubMed PMC
26. Ma S, Ji Z, Zhang B, et al. Spatial transcriptomic landscape unveils immunoglobin-associated senescence as a hallmark of aging. *Cell.* 2024;187:7025-44.e34. DOI
27. Yusri K, Kumar S, Fong S, Gruber J, Sorrentino V. Towards healthy longevity: comprehensive insights from molecular targets and biomarkers to biological clocks. *Int J Mol Sci.* 2024;25:6793. DOI PubMed PMC
28. Zhang L, Li T, Miao X, Ding L, Wang S, Wang Y. Aspirin suppresses *NFκB1* expression and inactivates cAMP signaling pathway to treat atherosclerosis. *Biofactors.* 2019;45:343-54. DOI
29. Van Der Wel T, Hilhorst R, Den Dulk H, et al. Chemical genetics strategy to profile kinase target engagement reveals role of FES in neutrophil phagocytosis. *Nat Commun.* 2020;11:3216. DOI PubMed PMC
30. Najm M, Azencott C, Playe B, Stoven V. Drug target identification with machine learning: how to choose negative examples. *Int J Mol Sci.* 2021;22:5118. DOI PubMed PMC
31. Hao Y, Quinnies K, Realubit R, Karan C, Tatonetti NP. Tissue-specific analysis of pharmacological pathways. *CPT Pharmacom Syst Pharma.* 2018;7:453-63. DOI PubMed PMC
32. Yu Z, Wang Y, Zhang Y, et al. PersADE: a database of personalized adverse drug events and their underlying molecular mechanisms. *Nucleic Acids Res.* 2026;54:D1556-68. DOI
33. Tong L, Ge Y, Yang S. Enhancing the confidence of potential targets enriched by similarity-centric models: the crucial role of the similarity threshold. *Front Pharmacol.* 2025;16:1574540. DOI PubMed PMC
34. De La Fuente J, Serrano G, Veleiro U, et al. Towards a more inductive world for drug repurposing approaches. *Nat Mach Intell.* 2025;7:495-508. DOI
35. Rask-Andersen M, Masuram S, Schiöth HB. The druggable genome: evaluation of drug targets in clinical trials suggests major shifts in molecular class and indication. *Annu Rev Pharmacol Toxicol.* 2014;54:9-26. DOI PubMed
36. Tanoli Z, Seemab U, Scherer A, Wennerberg K, Tang J, Vähä-Koskela M. Exploration of databases and methods supporting drug repurposing: a comprehensive survey. *Brief Bioinform.* 2021;22:1656-78. DOI PubMed PMC
37. Vanhaelen Q. Web-based tools for drug repurposing: successful examples of collaborative research. *Curr Med Chem.* 2020;28:181-95. DOI

Disclaimer/Publisher's Note: All statements, opinions, and data contained in this publication are solely those of the individual author(s) and contributor(s) and do not necessarily reflect those of OAE and/or the editor(s). OAE and/or the editor(s) disclaim any responsibility for harm to persons or property resulting from the use of any ideas, methods, instructions, or products mentioned in the content.



© The Author(s) 2026. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.