

EWAS Open Platform 2026: a deeply integrated resource for epigenome-wide association studies

Fei Yang^{1,2,†}, Zhuang Xiong^{3,†}, Wenting Zong^{1,2,†}, Demian Kong^{1,2,4,†}, Bixia Tang^{1,2}, Xupeng Chen³, Yaoke Wei^{1,2,4}, Xiangyu Yu^{1,2,4}, Yiran Zhang^{1,2,4}, Dong Zou^{1,2}, Zhang Zhang^{1,2,4,*}, Yiming Bao^{1,2,4,*}, Rujiao Li^{1,2,*}

¹National Genomics Data Center, China National Center for Bioinformation, Beijing100101, China

²Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing100101, China

³Interdisciplinary Institute for Medical Engineering, Fuzhou University, Fuzhou350002, China

⁴University of Chinese Academy of Sciences, Beijing100049, China

*To whom correspondence should be addressed. Email: lij@cnbc.ac.cn

Correspondence may also be addressed to Yiming Bao. Email: baoyim@cnbc.ac.cn

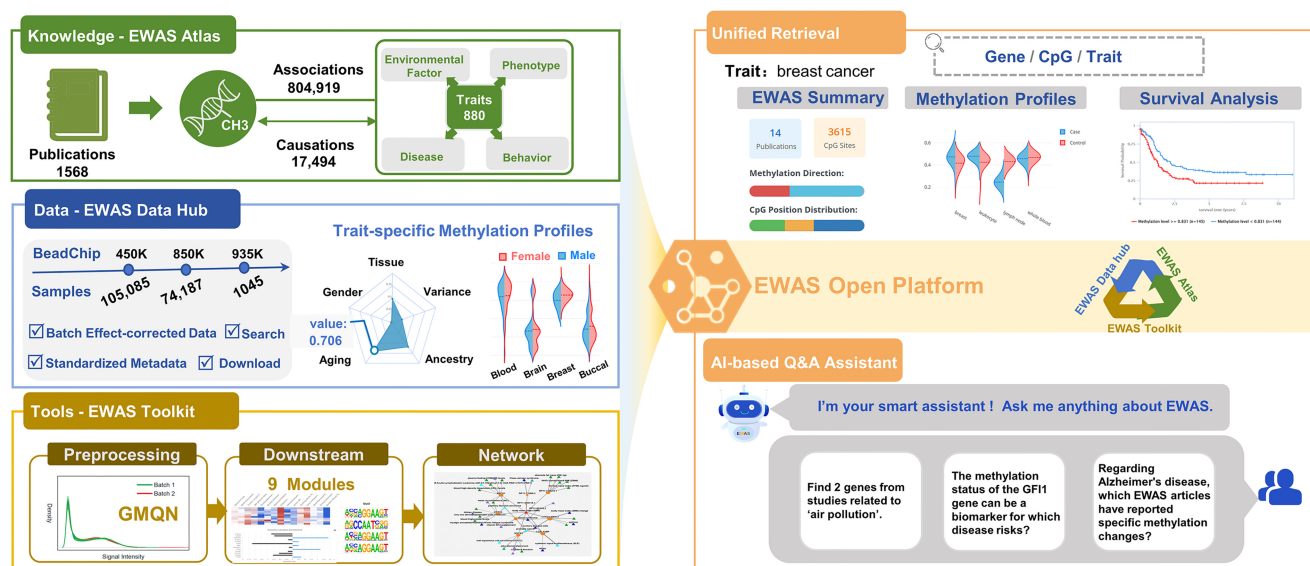
Correspondence may also be addressed to Zhang Zhang. Email: zhangzhang@cnbc.ac.cn

[†]The first four author should be regarded as Joint First Authors.

Abstract

Epigenome-wide association studies (EWAS) has become an indispensable approach for elucidating the epigenetic basis of complex traits. EWAS Open Platform (<https://ngdc.cnbc.ac.cn/ewas/>) includes three main components: EWAS Atlas (curated associations from publications), EWAS Data Hub (normalized DNA methylation array data), and EWAS Toolkit (one-stop analysis services). Here, we present a new release of EWAS Open Platform with the following significant updates and enhancements: (i) Expanded information: EWAS Atlas houses over 800 000 associations and incorporates an additional 17 000 curated causal relationships. EWAS Data Hub contains > 180 000 batch-corrected samples, which incorporate data from newly added Illumina MethylationEPIC v2.0 BeadChip (935K) as well as trait-specific methylation profiles. EWAS Toolkit now provides an online batch correction tool and an interactive epigenetic causal network. (ii) Enhanced interoperability: data, knowledge, and toolkit are fully interconnected through a unified retrieval, offering integrated summaries and visualization capabilities. (iii) Artificial intelligence (AI)-based service: the platform is newly equipped with an AI-assisted question-answering service, allowing users to interactively explore EWAS-related questions and generate tailored insights. Taken together, EWAS Open Platform has undergone a significant upgrade across data resources, analytical tools, and service functionalities, offering more advanced support for unraveling complex molecular mechanisms from an epigenomic perspective.

Graphical abstract



Received: September 15, 2025. Revised: October 9, 2025. Accepted: October 11, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Epigenome-wide association studies (EWAS) has emerged as a pivotal approach for elucidating the epigenetic basis of complex traits, including diseases [1–6]. The rapidly expanding volume of EWAS research and data enables in-depth integration and mining for epigenetic biomarker discovery and regulatory mechanism exploration. However, these large-scale data resources present several significant challenges: data are often dispersed and lack standardized formats, batch effects remain inadequately addressed, knowledge is fragmented with uncertain reliability, and analytical workflows are complex and cumbersome, together highlighting the urgent need for more systematic and efficient solutions [7]. To address this need, EWAS Open Platform (<https://ngdc.cncb.ac.cn/ewas/>) has been developed, including three components: EWAS Atlas [8], which focuses on the literature curation of EWAS associations; EWAS Data Hub [9], which is dedicated to the integration and normalization of large-scale methylation datasets; and EWAS Toolkit [10], which offers a suite of user-friendly, one-stop online tools for EWAS analysis. Ultimately, these components were unified into a comprehensive and scalable resource, with continuous updates and upgrades to support the full spectrum of EWAS [11–19].

To consistently provide up-to-date and comprehensive services, the updated EWAS Open Platform further strengthens the interoperability across data, knowledge, and toolkits (Table 1). EWAS Atlas adds a curated causal relationships module for more precise insights into epigenetic mechanisms. EWAS Data Hub expands data diversity, feature coverage, and trait-specific DNA methylation profiles. EWAS Toolkit now provides upstream batch correction for DNA methylation arrays, including the latest Illumina MethylationEPIC v2.0 BeadChip (935K), and downstream multi-omics causal networks. Beyond the three core components, the most transformative advancement in this release is the unification of data and knowledge through an integrated, visualized retrieval. Leveraging an artificial intelligence (AI)-based Q&A assistant, users can interactively explore EWAS-related queries and obtain personalized insights from the platform. This innovation positions the platform as a deeply integrated system that tightly links data access with biological interpretation, thereby offering researchers a more coherent, intelligent, and efficient resource for epigenetic investigations.

New features and updates

Knowledge—EWAS Atlas

To move beyond simple associations toward causal inference [20], the updated EWAS Atlas systematically integrates curated causal knowledge into its previously association-centric framework, thereby enabling more precise mechanistic insights and facilitating the discovery of potential intervention targets. This module incorporates causal relationships derived from Mendelian randomization studies [21–23], including 3836 methylation-to-trait, 45 trait-to-methylation, 10 555 methylation-to-expression, 1862 expression-to-trait, and 1193 trait-to-trait associations curated from publications. To facilitate efficient data exploration, the platform provides query functionality that allows searching by traits, methylation sites, and genes. Queries for diseases such as lung cancer reveal potential causal associations among methylation sites, gene expression events, and related traits (Fig. 1). This func-

tionality enables researchers to assess whether epigenetic variations are linked with specific traits and to investigate potential causal mechanisms underlying disease, thereby supporting the identification of biomarkers, intervention targets, and therapeutic candidates. Meanwhile, we have continuously tracked new publications on association studies, adding 189 142 trait-methylation associations and expanding the coverage to 262 new traits, including environmental factors, behaviors, phenotypes, and diseases. Overall, EWAS Atlas now hosts 806 860 high-quality EWAS associations and 17 494 causal links derived from 1568 publications, representing a >70% increase compared to the previous release. These span 248 tissues and cell types, 880 traits, and over 35 000 genes, providing an unprecedented breadth of coverage (Table 1).

Data—EWAS Data Hub

To keep pace with the latest technological advances in DNA methylation and enable high-resolution, trait-specific epigenomic analysis, EWAS Data Hub has now incorporated data from the newly released 935K array, alongside existing Illumina HumanMethylation450 BeadChip (450K) and Illumina MethylationEPIC v1.0 BeadChip (850K) arrays. It features harmonized methylation profiles for 180 317 samples from public sources, spanning 1413 tissues/cell types and 842 diseases. Notably, all newly incorporated data, including the 935K samples, underwent Gaussian mixture quantile normalization (GMQN)-based batch effect correction [24], ensuring consistency across samples for downstream analyses. In parallel, the 850K samples have undergone substantial growth, increasing from 20 923 to 74 156 samples, representing a >3.5-fold expansion. Leveraging this expanded data, we newly constructed comprehensive trait-specific methylation profiles at 898 918 CpG sites, characterizing distinct methylation patterns across tissues, age groups, sex, ethnicity, BMI, 76 cancer types, and 60 non-cancer diseases (Table 1).

Tools—EWAS Toolkit

To assist users in performing batch effect correction on their own methylation data and to enable seamless integration or comparative analyses with large-scale data in EWAS Data Hub, we have launched a one-click online batch effect correction tool based on GMQN [24], which supports data from not only 450K and 850K but also 935K arrays, thereby significantly enhancing its applicability and utility. The service enables users to submit data through multiple approaches, including uploading m/um files or directly inputting Gene Expression Omnibus (GEO) [25, 26] dataset accession numbers, thereby offering a more flexible and convenient solution for batch effect correction. Furthermore, to enable users to visualize disease-related molecular interaction networks and explore directed causal relationships, we incorporated a multi-layer user-interactive causal network using Cytoscape plugins, systematically integrating causal relationships at epigenomic, transcriptomic, and trait levels, thereby providing novel insights into molecular mechanisms underlying diseases. The network supports multidimensional queries based on traits, genes, and CpG sites, and allows users to flexibly adjust node ranges to generate directed causal maps centered on the queried entities. Taken together, EWAS Toolkit now offers 12 tools, delivering comprehensive and streamlined services that cover the entire analytical workflow, from upstream data

Table 1. Main updates of EWAS Open Platform

		EWAS Open Platform (2022)	EWAS Open Platform (2026)
Knowledge			
Publications		910	1568
Associations		617 018	806 860
Causations	Total	NA	17 494
	Methylation→trait	NA	3836
	Trait→methylation	NA	45
	Methylation→expression	NA	10 557
	Trait→trait	NA	1193
Traits		618	880
Studies		1437	1908
Cohorts		3382	3989
Data			
Samples	Total	115 852	180 317
	935K ^a	NA	1045
	850K ^b	20 923	74 187
	450K ^c	94 929	105 085
Tissues or cells		1118	1413
Diseases		528	842
Fields		242	296
Trait-specific methylation profiles		450K ^c	450K ^c , 850K ^b
Toolkit			
Causation network		NA	Yes
Batch effect correction—GMQN ^d		NA	Yes
Unified retrieval		NA	Yes
AI- based Q&A assistant		NA	Yes

^aIllumina MethylationEPIC v2.0 BeadChip.^bIllumina MethylationEPIC v1.0 BeadChip.^cIllumina HumanMethylation450 BeadChip.^dGaussian mixture quantile normalization.

preprocessing to downstream functional annotation and exploration.

Unified retrieval

To address the fragmentation of data, knowledge, and toolkit on EWAS Open Platform and to enable efficient, integrated access, we developed a seamless cross-module unified retrieval that combines quantitative summaries with interactive visualizations. This interface leverages a unified application programming interface (API), integrates heterogeneous data through semantic parsing, and employs standardized data modeling, ensuring consistency, interoperability, and robust query support across diverse data types. Users can query by traits, genes, or CpG probe IDs, with results delivered through concise summaries and intuitive visual outputs. For instance, a query for “breast cancer” returns comprehensive quantitative summaries, including the number and categories of methylation sites and genes reported in publications as associated with breast cancer within EWAS Atlas, along with corresponding number of sample from EWAS Data Hub. Additional outputs include the proportions of hyper- and hypomethylated sites, the genomic distribution of CpG sites, and a word cloud highlighting the most frequently reported differentially methylated genes (Fig. 2A). Users can further examine detailed information on representative methylation sites and genes associated with breast cancer (Fig. 2B and C). Moreover, the platform illustrates the methylation patterns of these genes in breast cancer and their associations with patient survival, based on analyses from EWAS Data Hub (Fig. 2D), and presents the epigenetic association network related to breast cancer (Fig. 2E). Embedded links provide seamless navigation and cross-module exploration of related content.

AI-based Q&A assistant

To enable researchers to access epigenetic information on complex traits including diseases in a more flexible and intuitive manner, we developed an intelligent agent-based question-answering assistant within EWAS Open Platform, using the LangChain and LangGraph frameworks [27, 28] along with Vue3 [29] (Fig. 3A). Technically, the assistant implements a multistage text-to-SQL pipeline, including intelligent entity recognition, context-aware SQL generation, data querying and structuring, and enrichment and secondary processing, as well as final delivery to the frontend. Retrieved data are displayed in the user interface for in-depth exploration and analysis. Meanwhile, the processed results are passed to a large language model, which produces clear, accurate, and informative natural language answers, effectively transforming unstructured queries into actionable epigenetic insights through intelligent reasoning.

The assistant supports conversational queries in natural language and automatically returns structured answers derived from integrated EWAS resources, covering disease associations, methylation patterns, and multidimensional factors such as age, sex, ethnicity, and tissue. For example, when asked “The methylation status of the *GF11* gene can be a biomarker for which disease risks?,” the assistant compiles a panel summarizing diseases associated with *GF11* and its involvement in biological processes. In addition, the assistant provides a radar chart of *GF11*–trait associations and visualizes methylation sites and their characteristics (Fig. 3B). By linking unstructured queries to comprehensive EWAS resources, the AI-based assistant offers researchers, including those who are not specialists, a more direct and flexible way to explore complex epigenetic information and insights. It organizes and presents these insights into intuitive visualizations and structured

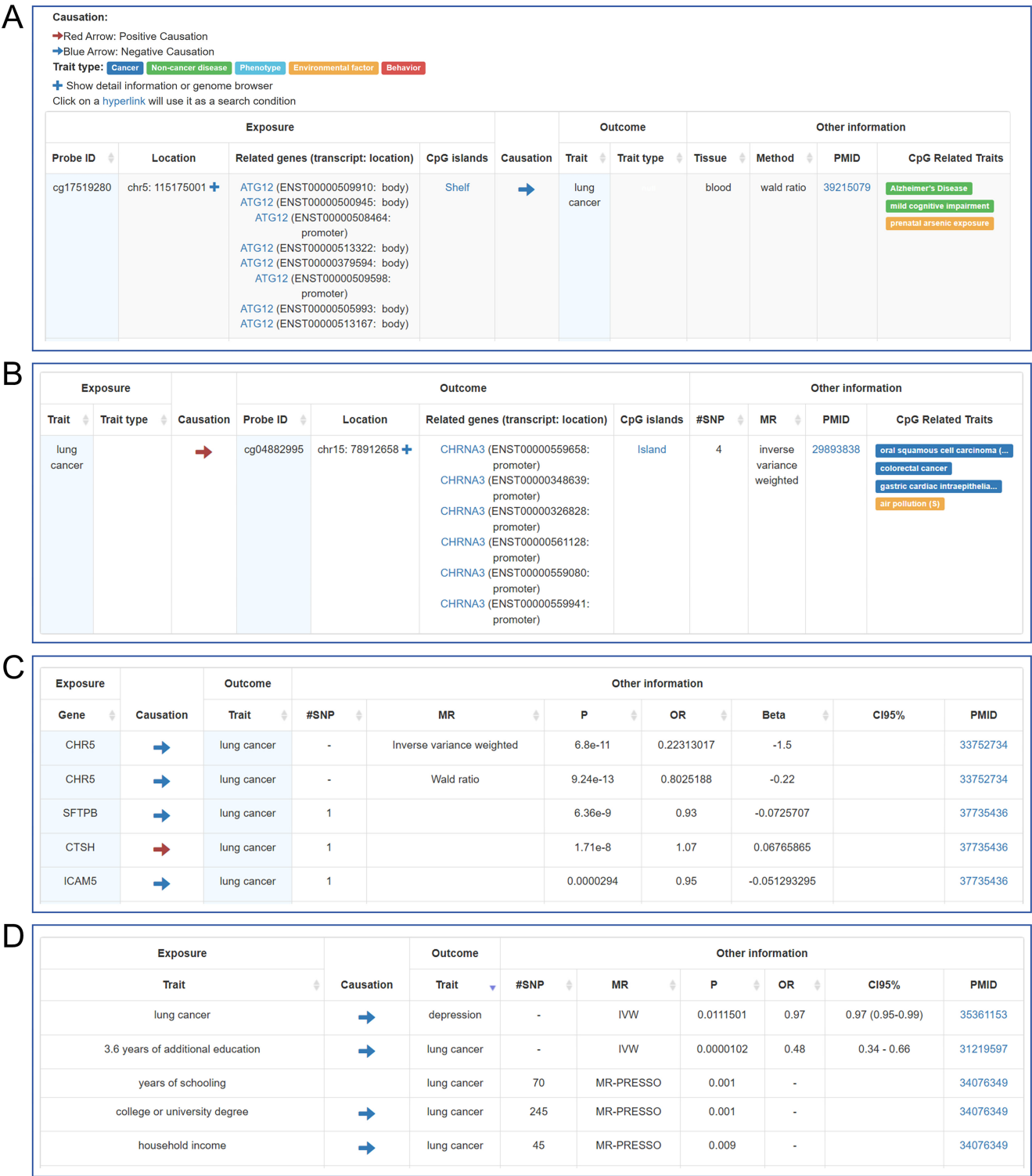


Figure 1. Retrieval results from the causal knowledge module using lung cancer as an example. **(A)** Causal relationships from DNA methylation to lung cancer. **(B)** Causal relationships from lung cancer to DNA methylation. **(C)** Causal relationships from gene expression to lung cancer. **(D)** Causal relationships between lung cancer and other traits.

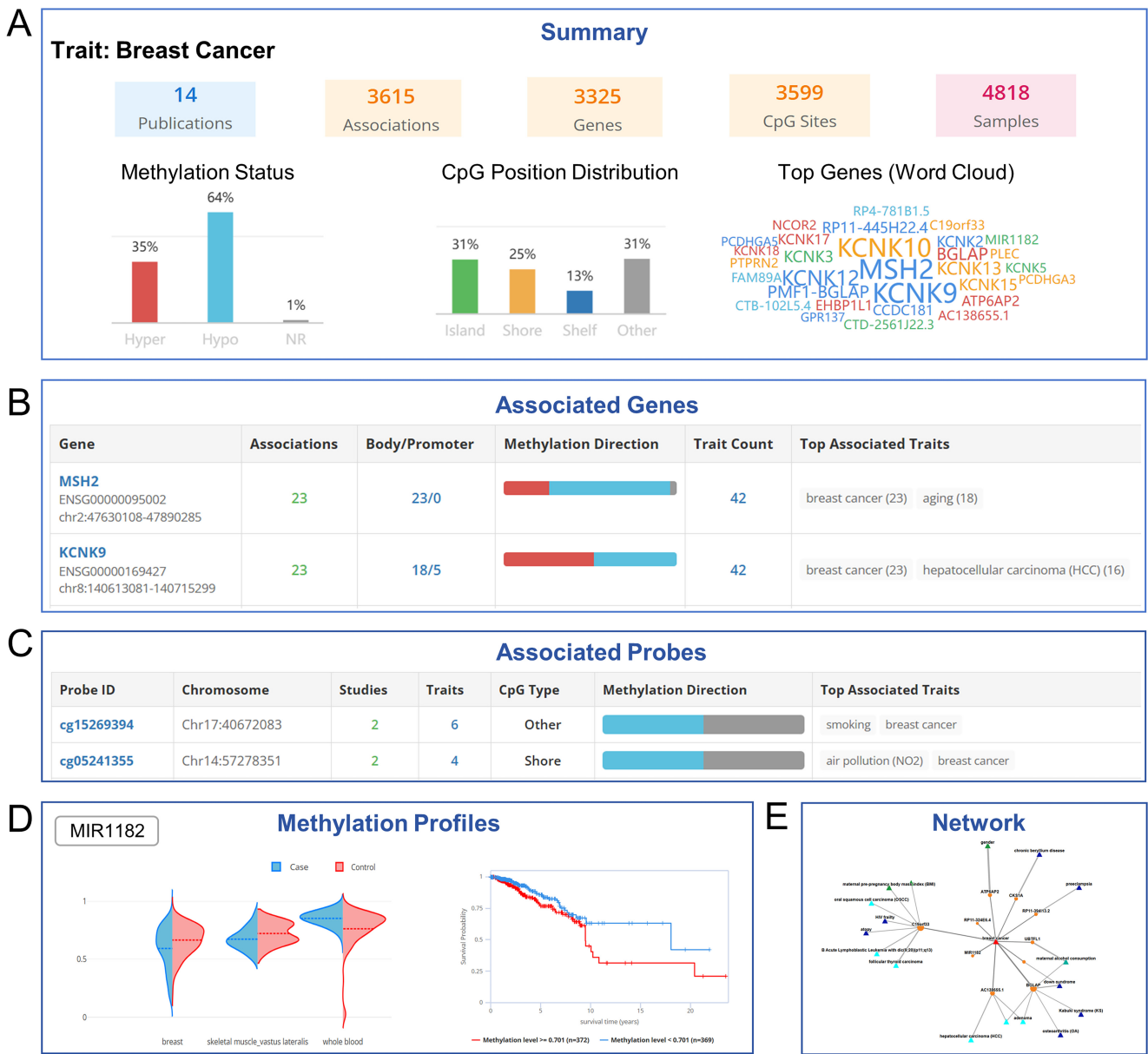


Figure 2. Unified retrieval results from EWAS Open Platform using breast cancer as an example. **(A)** Overview of EWAS publications, datasets, and associated methylation sites related to breast cancer. **(B)** Characteristics of representative methylated genes associated with breast cancer. **(C)** Features of representative methylation sites associated with breast cancer. **(D)** Differential methylation patterns between breast cancer and normal samples across different tissue types and their association with patient survival. **(E)** Epigenetic association network related to breast cancer.

summaries, facilitating interpretation, supporting the identification of potential associations, and leveraging intelligent reasoning to guide new directions in epigenetic research.

Discussion and future developments

EWAS Open Platform, as one of the essential resources within National Genomics Data Center, China National Center for Bioinformation [30, 31], continues to provide researchers with an integrated suite of data, knowledge, and toolkit to support EWAS. Furthermore, the high-quality datasets available through this platform have already facilitated the training of several AI models [32–34]. This update delivers significant advancements across multiple dimensions, including the targeted expansion of causal knowledge, the enhancement of data profiles, the integration of upstream and downstream analytical

tools, AI-based query capabilities, and the implementation of comprehensive, interactive data-knowledge retrieval functionalities. Collectively, these improvements brings EWAS Open Platform closer to its vision of serving as a robust and integrative resource for advancing epigenomic research.

Looking ahead, we plan to further embed AI within the epigenetic knowledge curation framework and to establish more automated pipelines for data analysis, thereby enhancing the efficiency and rigor of knowledge discovery. Concurrently, EWAS Open Platform will not only strengthen its connections with resources, such as MethBank [35–37], scMethBank [5], and DiseaseMeth [38], but also enhance data quality, enrich annotation content, and extend its scope beyond DNA methylation to encompass a broader spectrum of epigenetic associations [39–46], thereby expanding the depth and breadth of integrative studies and empowering AI-driven

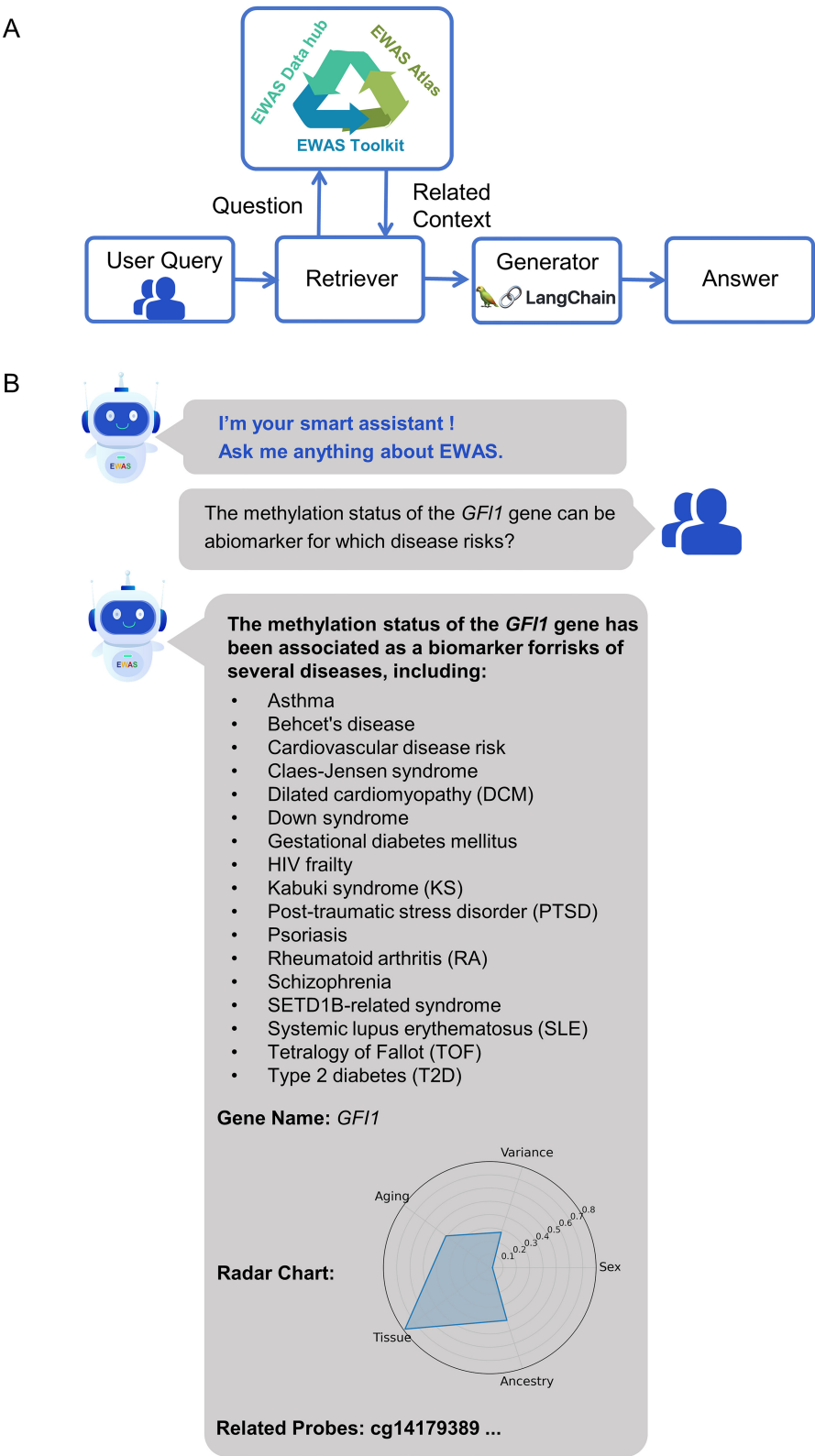


Figure 3. The AI-based question-answering assistant of EWAS Open Platform. **(A)** The architecture of AI-assisted Q&A assistant. **(B)** Example output for the query: "The methylation status of the *GFI1* gene can be a biomarker for which disease risks?"

analyses. In terms of implementation, we aim to develop end-to-end, sample-selectable workflows alongside user-friendly online tools for differential methylation analysis and functional annotation. In parallel, the integration of EWAS with genome-wide association study (GWAS) and transcriptome-wide association study (TWAS) will be strengthened to foster multi-omics synergy and to address critical scientific challenges in population health research [12, 47–50]. Collectively, these improvements will provide the research community with a versatile, interoperable, and sustainable platform, further enabling AI applications, accelerating epigenomic discovery, and elucidating the complex mechanisms underlying human health and disease.

Acknowledgements

We wish to express our heartfelt gratitude to Mengwei Li of Cancer Research UK Scotland Institute for his continuous assistance and technical support with EWAS Open Platform. We also sincerely thank the numerous users who have reported bugs and provided constructive suggestions, which have been invaluable for the continuous improvement of the platform.

Author contributions: Fei Yang (Conceptualization [lead], Data curation [lead], Formal analysis [lead], Funding acquisition [supporting], Software [equal], Visualization [equal], Writing—original draft [lead]), Zhuang Xiong (Conceptualization [supporting], Data curation [supporting], Formal analysis [equal], Software [equal], Writing—review & editing [equal]), Wenting Zong (Conceptualization [supporting], Formal analysis [equal], Software [equal], Visualization [equal], Writing—review & editing [equal]), Demian Kong (Methodology [supporting], Software [equal], Visualization [equal], Writing—review & editing [supporting]), Bixia Tang (Methodology [equal], Software [supporting], Supervision [supporting], Writing—review & editing [supporting]), Xupeng Chen (Data curation [supporting], Formal analysis [equal]), Yaoke Wei (Data curation [equal]), Xianguyu Yu (Data curation [equal]), Yiran Zhang (Data curation [equal]), Dong Zou (Methodology [supporting], Software [supporting], Visualization [supporting]), Zhang Zhang (Conceptualization [supporting], Funding acquisition [supporting], Project administration [equal], Supervision [lead], Writing—review & editing [lead]), Yiming Bao (Conceptualization [lead], Funding acquisition [lead], Project administration [lead], Supervision [lead], Writing—review & editing [lead]) and Ruijiao Li (Conceptualization [lead], Funding acquisition [supporting], Investigation [equal], Project administration [lead], Supervision [lead], Writing—original draft [supporting], Writing—review & editing [lead]).

Conflict of interest

All authors have confirmed that there are no conflicts of interest to disclose.

Funding

The Informatization Plan of Chinese Academy of Sciences (CAS-WX2024GC-0602), Strategic Priority Research Program of the Chinese Academy of Sciences (XDA0460202, XDA0460401), National Natural Science Foundation of China (32300468, 32030021, 32400472), Genomics Data Center Operation and Maintenance of Chinese Academy of

Sciences (CAS-WX2022SDC-XK05), The Open Biodiversity and Health Big Data Programme of IUBS. Funding to pay the Open Access publication charges for this article was provided by the Informatization Plan of Chinese Academy of Sciences (CAS-WX2024GC-0602).

Data availability

EWAS Open Platform is a comprehensive resource for EWAS. All data, knowledge, and tools are freely available at <https://ngdc.cncb.ac.cn/ewas/>. Any queries, comments, and suggestions on EWAS Open Platform can be provided by email via ewas-user@big.ac.cn.

References

1. Wei S, Tao J, Xu J *et al*. Ten years of EWAS. *Adv Sci* 2021;8:e2100727. <https://doi.org/10.1002/adv.202100727>
2. Smith ZD, Hetzel S, Meissner A. DNA methylation in mammalian development and disease. *Nat Rev Genet* 2025;26:7–30. <https://doi.org/10.1038/s41576-024-00760-8>
3. Flanagan JM. Epigenome-wide association studies (EWAS): past, present, and future. *Methods Mol Biol* 2015;1238:51–63.
4. Osorio J. Obesity. Looking at the epigenetic link between obesity and its consequences—the promise of EWAS. *Nat Rev Endocrinol* 2014;10:249. <https://doi.org/10.1038/nrendo.2014.42>
5. Zong W, Kang H, Xiong Z *et al*. scMethBank: a database for single-cell whole genome DNA methylation maps. *Nucleic Acids Res* 2022;50:D380–6. <https://doi.org/10.1093/nar/gkab833>
6. Jiang Y, Qu M, Jiang M *et al*. MethylGenotyper: accurate estimation of SNP genotypes and genetic relatedness from DNA methylation data. *Genomics Proteomics Bioinformatics* 2024;22:qzae044. <https://doi.org/10.1093/gpbjnl/qzae044>
7. Breeze CE, Wong JYY, Beck S *et al*. Diversity in EWAS: current state, challenges, and solutions. *Genome Med* 2022;14:71. <https://doi.org/10.1186/s13073-022-01065-3>
8. Li M, Zou D, Li Z *et al*. EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res* 2019;47:D983–8. <https://doi.org/10.1093/nar/gky1027>
9. Xiong Z, Li M, Yang F *et al*. EWAS Data Hub: a resource of DNA methylation array data and metadata. *Nucleic Acids Res* 2020;48:D890–5. <https://doi.org/10.1093/nar/gkz840>
10. Xiong Z, Yang F, Li M *et al*. EWAS Open Platform: integrated data, knowledge and toolkit for epigenome-wide association study. *Nucleic Acids Res* 2022;50:D1004–9. <https://doi.org/10.1093/nar/gkab972>
11. Peng Q, Liu X, Li W *et al*. Analysis of blood methylation quantitative trait loci in East Asians reveals ancestry-specific impacts on complex traits. *Nat Genet* 2024;56:846–60. <https://doi.org/10.1038/s41588-023-01494-9>
12. Shao M, Chen K, Zhang S *et al*. Multiome-wide association studies: novel approaches for understanding diseases. *Genomics Proteomics Bioinformatics* 2024;22:qzae077. <https://doi.org/10.1093/gpbjnl/qzae077>
13. Dabrowski JK, Yang EJ, Crofts SJC *et al*. Probabilistic inference of epigenetic age acceleration from cellular dynamics. *Nat Aging* 2024;4:1493–507. <https://doi.org/10.1038/s43587-024-00700-5>
14. Crofts SJC, Latorre-Crespo E, Chandra T. DNA methylation rates scale with maximum lifespan across mammals. *Nat Aging* 2024;4:27–32. <https://doi.org/10.1038/s43587-023-00535-6>
15. Smyth LJ, Dahlstrom EH, Syreeni A *et al*. Epigenome-wide meta-analysis identifies DNA methylation biomarkers associated with diabetic kidney disease. *Nat Commun* 2022;13:7891. <https://doi.org/10.1038/s41467-022-34963-6>
16. van Dongen J, Gordon SD, McRae AF *et al*. Identical twins carry a persistent epigenetic signature of early genome programming. *Nat Commun* 2021;12:5618. <https://doi.org/10.1038/s41467-021-25583-7>

17. Wielscher M, Mandaviya PR, Kuehnel B *et al.* DNA methylation signature of chronic low-grade inflammation and its role in cardio-respiratory diseases. *Nat Commun* 2022;13:2408. <https://doi.org/10.1038/s41467-022-29792-6>
18. Shore CJ, Villicana S, El-Sayed Moustafa JS *et al.* Genetic effects on the skin methylome in healthy older twins. *Am J Hum Genet* 2024;111:1932–52. <https://doi.org/10.1016/j.ajhg.2024.07.010>
19. Ouni M, Kovac L, Gancheva S *et al.* Novel markers and networks related to restored skeletal muscle transcriptome after bariatric surgery. *Obesity* 2024;32:363–75. <https://doi.org/10.1002/oby.23954>
20. Pingault JB, O'Reilly PF, Schoeler T *et al.* Using genetic data to strengthen causal inference in observational research. *Nat Rev Genet* 2018;19:566–80. <https://doi.org/10.1038/s41576-018-0020-3>
21. Birney E. Mendelian randomization. *Cold Spring Harb Perspect Med* 2022;12:a041302.
22. Grau-Perez M, Agha G, Pang Y *et al.* Mendelian randomization and the environmental epigenetics of health: a systematic review. *Curr Environ Health Rep* 2019;6:38–51. <https://doi.org/10.1007/s40572-019-0226-3>
23. Juvinao-Quintero DL, Hivert MF, Sharp GC *et al.* DNA methylation and type 2 diabetes: the use of Mendelian randomization to assess causality. *Curr Genet Med Rep* 2019;7:191–207. <https://doi.org/10.1007/s40142-019-00176-5>
24. Xiong Z, Li M, Ma Y *et al.* GMQN: a reference-based method for correcting batch effects and probe bias in HumanMethylation BeadChip. *Front. Genet.* 2021;12:810985. <https://doi.org/10.3389/fgene.2021.810985>
25. Barrett T, Troup DB, Wilhite SE *et al.* NCBI GEO: archive for functional genomics data sets—10 years on. *Nucleic Acids Research* 2011;39:D1005–10. <https://doi.org/10.1093/nar/gkq1184>
26. Barrett T, Wilhite SE, Ledoux P *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res* 2013;41:D991–5. <https://doi.org/10.1093/nar/gks1193>
27. Workman AD, Rathi VK, Lerner DK *et al.* Utility of a LangChain and OpenAI GPT-powered chatbot based on the international consensus statement on allergy and rhinology: rhinosinusitis. *Int Forum Allergy Rhinol* 2024;14:1101–9. <https://doi.org/10.1002/alr.23310>
28. Topsakal O, Akinci TC. Creating large language model applications utilizing LangChain: a primer on developing LLM apps fast. *ICAENS* 2023;1:1050–6. <https://doi.org/10.59287/icaens.1127>
29. Li C, Zhang SY, Na JS *et al.* Design and implementation of project management system based on Vue3.0 framework. In: *2nd International Conference on Algorithms, High Performance Computing and Artificial Intelligence (AHPCAII)*. IEEE, 2022, 347–53. <https://doi.org/10.1109/AHPCAII57455.2022.10087744>
30. CNCB-NGDC Members and Partners. Database resources of the National Genomics Data Center, China National Center for Bioinformatics in 2025. *Nucleic Acids Res* 2025;53:D30–44. <https://doi.org/10.1093/nar/gkae978>
31. CNCB-NGDC Members and Partners. Database resources of the National Genomics Data Center, China National Center for Bioinformatics in 2024. *Nucleic Acids Res* 2024;52:D18–32. <https://doi.org/10.1093/nar/gkad1078>
32. Urban A, Sidorenko D, Zagirova D *et al.* Precious1GPT: multimodal transformer-based transfer learning for aging clock development and feature importance analysis for aging and age-related disease target discovery. *Aging* 2023;15:4649–66.
33. Zhao M, Cai M, Lei F *et al.* AI-driven feature selection and epigenetic pattern analysis: a screening strategy of CpGs validated by pyrosequencing for body fluid identification. *Forensic Sci Int* 2025;367:112339. <https://doi.org/10.1016/j.forsciint.2024.112339>
34. Ying K, Song J, Cui H *et al.* MethylGPT: a foundation model for the DNA methylome. *bioRxiv*, <https://doi.org/10.1101/2024.10.30.621013>, 3 November 2024, preprint: not peer reviewed.
35. Zhang M, Zong W, Zou D *et al.* MethBank 4.0: an updated database of DNA methylation across a variety of species. *Nucleic Acids Res* 2023;51:D208–16. <https://doi.org/10.1093/nar/gkac969>
36. Zou D, Sun S, Li R *et al.* MethBank: a database integrating next-generation sequencing single-base-resolution DNA methylation programming data. *Nucleic Acids Res* 2015;43:D54–8. <https://doi.org/10.1093/nar/gku920>
37. Li R, Liang F, Li M *et al.* MethBank 3.0: a database of DNA methylomes across a variety of species. *Nucleic Acids Res* 2018;46:D288–95. <https://doi.org/10.1093/nar/gkx1139>
38. Xing J, Zhai R, Wang C *et al.* DiseaseMeth version 3.0: a major expansion and update of the human disease methylation database. *Nucleic Acids Res* 2022;50:D1208–15. <https://doi.org/10.1093/nar/gkab1088>
39. Kiri S, Ryba T. Cancer, metastasis, and the epigenome. *Mol Cancer* 2024;23:154. <https://doi.org/10.1186/s12943-024-02069-w>
40. Burton A, Torres-Padilla ME. Epigenome dynamics in early mammalian embryogenesis. *Nat Rev Genet* 2025;26:587–603. <https://doi.org/10.1038/s41576-025-00831-4>
41. Terekhanova NV, Karpova A, Liang WW *et al.* Epigenetic regulation during cancer transitions across 11 tumour types. *Nature* 2023;623:432–41. <https://doi.org/10.1038/s41586-023-06682-5>
42. Gershman A, Sauria MEG, Guitart X *et al.* Epigenetic patterns in a complete human genome. *Science* 2022;376:eabj5089. <https://doi.org/10.1126/science.abj5089>
43. Huang S, Han Y, Liu Y *et al.* Super-enhancer-mediated circRNAs exhibit high splicing circularization diversity and transcriptional activity. *Nucleic Acids Res* 2025;53:gkaf505. <https://doi.org/10.1093/nar/gkaf505>
44. Lu Y, Yang L, Feng Q *et al.* RNA 5-methylcytosine modification: regulatory molecules, biological functions, and human diseases. *Genomics Proteomics Bioinformatics* 2024;22:qzae063. <https://doi.org/10.1093/gpbjnl/qzae063>
45. Song B, Huang D, Zhang Y *et al.* m6A-TSHub: unveiling the context-specific m⁶A methylation and m⁶A-affecting mutations in 23 human tissues. *Genomics Proteomics Bioinformatics* 2023;21:678–94. <https://doi.org/10.1016/j.gpb.2022.09.001>
46. Karlow JA, Pehrsson EC, Xing X *et al.* Non-small cell lung cancer epigenomes exhibit altered DNA methylation in smokers and never-smokers. *Genomics Proteomics Bioinformatics* 2023;21:991–1013. <https://doi.org/10.1016/j.gpb.2023.03.006>
47. Li Y, van den Berg EH, Kurilshikov A *et al.* Genome-wide studies reveal genetic risk factors for hepatic fat content. *Genomics Proteomics Bioinformatics* 2024;22:qzae031. <https://doi.org/10.1093/gpbjnl/qzae031>
48. Chai Y, Zhang R, Jia S *et al.* Toti: an integrated multi-omics database to decipher the epigenetic regulation of gene expression in totipotent stem cells. *Stem Cell Rep* 2025;20:102605. <https://doi.org/10.1016/j.stemcr.2025.102605>
49. Li H, Zhao W, Yang F *et al.* Immunosenescence inventory—a multi-omics database for immune aging research. *Nucleic Acids Res* 2025;53:D1047–54. <https://doi.org/10.1093/nar/gkae1102>
50. Ding J, Blencowe M, Nghiem T *et al.* Mergeomics 2.0: a web server for multi-omics data integration to elucidate disease networks and predict therapeutics. *Nucleic Acids Res* 2021;49:W375–87. <https://doi.org/10.1093/nar/gkab405>