BMC Genomics

DATABASE

Open Access

MeDReaders 2.0: an updated database for modified DNA readers

Hongfei Li[1,3], Murong Zhou[2], Ximei Luo[1,3*] and Guohua Wang[2*]

Abstract

Here, we present a significant update to the MeDReaders database and webserver. The new MeDReaders 2.0 covers interactions between transcription factors (TFs) and DNA cytosine modifications that arise during DNA methylation and active demethylation, including 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). Based on TF–modified-DNA interactions curated from the literature, we extended the original MeDReaders dataset by adding 257 TF–5mC–DNA, 160 TF–5hmC–DNA, 220 TF–5fC–DNA, and 107 TF–5caC–DNA interactions. Using in silico approaches, we integrated whole-genome bisulfite sequencing (WGBS) and TF ChIP-seq datasets from the ENCODE and NCBI GEO databases to predict 1,363 TF–5mC–DNA interactions across six human cell lines and one mouse cell line. MeDReaders 2.0 provides a more comprehensive resource for studying how DNA cytosine modifications regulate TF–DNA interactions and for designing related experiments, such as those on dynamic DNA methylation and active demethylation. The updated website is accessible at https://bioinfor.nefu.edu.cn/MeDReaders2.0/.

Keyword Transcription factor, DNA methylation, Cytosine modification

Introduction

The classical view of DNA epigenetic modifications is that they usually prevent transcription factors (TFs) from binding to DNA [1–3]. TFs are traditionally considered to bind to open chromatin regions lacking DNA modifications. Modifications on cis-regulatory elements, such as 5-methylcytosine (5mC), can prevent TF binding or recruitment [4–7]. Only a limited number of proteins possess a specific mCpG-binding domain (MBD) that enables them to recognize and bind 5-methylcytosine-modified DNA (5mC-DNA) in a sequence-independent manner [8–10]. However, over the past 10 years, many studies have reported that TFs can interact with 5mC-DNA and that their selective binding to target DNA is affected by DNA epigenetic modifications [11–13]. Further in vivo and in vitro experiments have revealed the widespread existence of TFs binding to 5mC-DNA. DNA modifications, such as 5mC, can indeed alter the interactions between TFs and DNA [14, 15]. They can affect the DNA binding motifs as well as the binding affinity of TFs. Some transcription factors, like KLF4 [16], Kaiso [17], ZFP57 [18], and CEBPα [19], show a high affinity for distinct 5mC-DNA sequences. It was also found that the effect of epigenetic modifications on TF-DNA affinity can be binary. With the development of technology, additional epigenetic modifications beyond methylation have been explored. Various DNA modifications can have differing impacts on the affinity of TF-DNA

*Correspondence:
Ximei Luo
luoximei@uestc.edu.cn
Guohua Wang
ghwang@nefu.edu.cn
[1]Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China
[2]College of Computer and Control Engineering, Northeast Forestry University, Harbin 150040, China
[3]Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou 324003, China

interactions [20–23]. For example, the OVOL2-DNA interaction is significantly suppressed by modifications such as 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). On the other hand, carboxylation actually enhances the HMBOX1-DNA interaction. The identification and elucidation of TFs involved in methylation-mediated biological processes are critical steps in understanding the underlying mechanisms. By studying these TFs, researchers can gain valuable insights into how 5mC-DNA influences gene regulation and cellular processes [24]. For example, 5mC-DNA imbalances can disrupt TF binding, leading to dysregulated gene expression and thereby increasing susceptibility to diseases [24, 25]. Ultimately, deciphering these intricate mechanisms could lead to novel therapeutic strategies and interventions to improve human health [25].

To systematically decode these disease-relevant regulatory networks, advanced high-throughput technologies have been rapidly developed and widely applied. With the development of next-generation sequencing technologies and the accumulation of data, especially the application of systematic evolution of ligands by exponential enrichment (SELEX) [26], ChIP-BS-seq [27] and Digital Affinity Profiling via Proximity Ligation (DAPPL) [28], more epigenetic effects on TF-DNA interaction have been discovered. This explosion of data necessitates a more comprehensive and updated resource. Therefore, we upgraded MeDReaders 1.0 to MeDReaders 2.0 in two major ways [29]. First, leveraging newly released TF ChIP-seq datasets and matched whole-genome bisulfite sequencing (WGBS) data from public resources (e.g., ENCODE and GEO), we expanded our in silico pipeline to predict TF–5mC–DNA interactions in additional cell lines and increased the number of predicted TF–5mC–DNA interaction records in the originally covered core cell lines. Second, we extended the database to include TF interactions with other cytosine DNA modifications beyond 5mC—such as 5hmC, 5fC, and 5caC—by manually curating experimentally supported interactions from the literature, including those identified by mass spectrometry (MS) and DAPPL. All updates are freely available at https://bioinfor.nefu.edu.cn/MeDReaders2.0/.

Materials and methods

Data sources

In MeDReaders 2.0, the updated TF–epigenetically modified DNA interaction data come from two sources. First, we curated experimentally supported interactions from the literature. In MeDReaders 1.0, only TF–5mC–DNA interactions were collected from the literature [16–19, 26, 30–33], lacking TF interactions with other cytosine modifications; therefore, in MeDReaders 2.0, we newly curated 257 TF–5mC–DNA, 160 TF–5hmC–DNA, 220 TF–5fC–DNA, and 107 TF–5caC–DNA interactions from the literature [28, 34]. This diversity of cytosine modifications enables a more comprehensive explanation of the specificity of TF interactions during DNA methylation and active demethylation. Second, we updated the TF–5mC–DNA interactions predicted by an in silico pipeline. MeDReaders 1.0 predicted only 292 interactions. With updates in the ENCODE and GEO databases and the availability of more TF ChIP-seq and WGBS datasets, MeDReaders 2.0 expanded the in silico predictions to 1,363 TF–5mC–DNA interactions across six human cell lines (GM12878: 146; H1-hESC: 58; HeLa: 54; HepG2: 584; K562: 457; SK-N-SH: 59) and one mouse cell line (E14: 5).

Database implementation

Based on in silico predicted TF-5mC-DNA interactions, our study integrated TF ChIP-seq and WGBS data to investigate the interplay between TF binding and DNA methylation. We used ENCODE-processed ChIP-seq and WGBS datasets [35–38]. TF binding regions identified via ChIP-seq were comprehensively analyzed and categorized based on their corresponding methylation status. Motif discovery and scanning were performed using HOMER software, which generated a Position Probability Matrix (PPM) and subsequently converted it into a Position Weight Matrix (PWM) for sequence scoring. This PWM was employed in a dual-strand sliding window scan across TF peaks to pinpoint the best-matching binding sequences. These best-hit sequences were then aligned with WGBS-derived CpG coverage and methylation proportion data to extract specific CpG positions and their methylation levels. To construct methylation-aware motif logos, cytosines (C) at CpG sites exhibiting high methylation levels were explicitly replaced with the character 'E'. This sequence-level transformation created a 5-letter sequence format (A/C/G/T/E). Finally, position-specific frequencies were calculated from these sequences, and customized motif logos were plotted to visually represent highly methylated cytosines using the 'E' symbol.

The Python-based Flask framework is utilized for website development due to its simplicity in design philosophy and efficient development model. The web pages are rendered using the Jinja2 template engine, while the front-end and back-end communicate asynchronously through AJAX technology. The Semantic UI framework continues to be employed for designing and developing the front-end interface.

Result

Usage and access

In the updated version, we have improved the user-friendliness of the interface, making it convenient for

users to browse, search, and download TF binding information under various DNA epigenetic modifications. In addition to expanding our coverage of 5mC-DNA data, MeDReaders 2.0 now includes other cytosine modifications. We have visualized these different types of modifications to help users easily distinguish between them (Fig. 1A). Furthermore, the increasing availability of TF ChIP-seq data has enabled the identification of more TF-DNA interactions. The newly added TF–5mC–DNA interactions, obtained via our in silico pipeline, are primarily derived from the K562, HepG2, and GM12878 cell lines. Remarkably, the number of predicted interactions in these three cell lines has increased notably compared to the original version [29] (Fig. 1B). Additionally, the HeLa and SK-N-SH cell lines are newly covered in this update.

In MeDReaders 2.0, in silico predictions of TF–5mC–DNA interactions reveal both conserved and cell-type-specific patterns. As shown in Fig. 1C, a subset of TFs displays conserved interactions, with 13 TFs shared across all six human cell lines. In contrast, many TF–5mC–DNA interactions are highly cell-type-specific. For example, out of the 584 predicted TF–5mC–DNA interactions in the HepG2 cell line, 341 involve TFs that are observed exclusively in this cell type. These patterns provide an opportunity to investigate the cell-type specificity of 5mC-modified DNA in shaping TF binding and to gain insights into transcriptional regulation and cellular specificity.

Cytosine modification derivatives (such as 5hmC, 5fC, and 5caC) play a crucial role in the dynamic regulation of DNA methylation. Recognizing their importance, our database has newly curated literature-derived interactions between TFs and four distinct types of cytosine-modified DNA. As illustrated in Fig. 1D, while a substantial number of TFs can interact with multiple types of modified DNA, many TFs also exhibit a remarkable degree of modification specificity. For instance, out of the 220 identified TF–5fC–DNA interactions, 113 TFs interact exclusively with 5fC. Similarly, among the 107 TF–5caC–DNA interactions, 32 TFs are uniquely bound to 5caC. These findings provide a valuable foundation

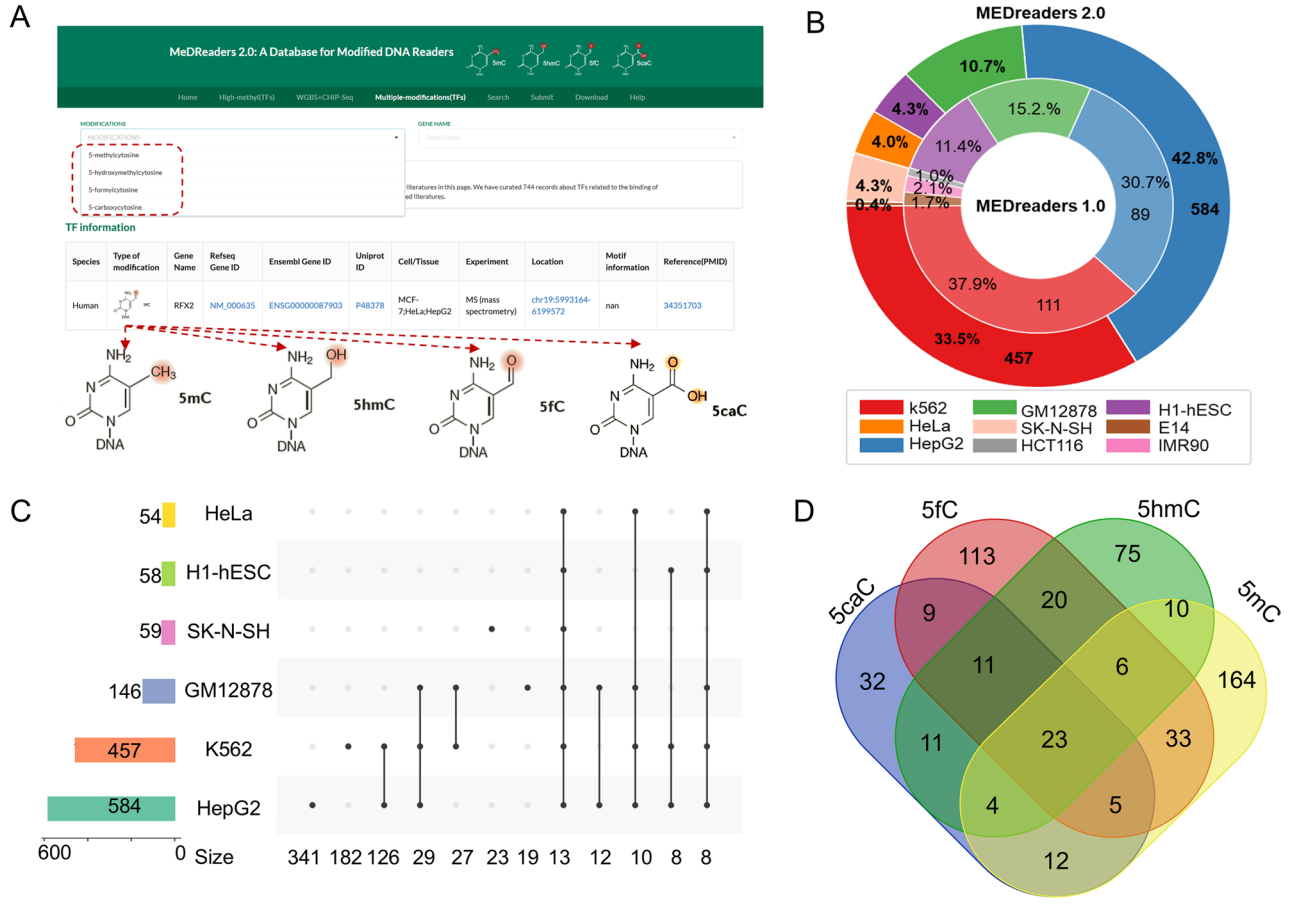

Fig. 1 Overview of data expansion and visualization in MeDReaders 2.0. **A** Web interface for TF-modified-cytosine-DNA interactions (5mC, 5hmC, 5fC, 5caC) curated from the literature. **B** Quantitative distribution of TF-5mC-DNA interactions predicted via the in silico pipeline across different cell lines in MeDReaders 1.0 and 2.0. **C** Intersection of TFs involved in in silico predicted TF-5mC-DNA interactions across six human cell lines. **D** Venn diagram of literature-curated TF-DNA interactions across four types of cytosine modifications

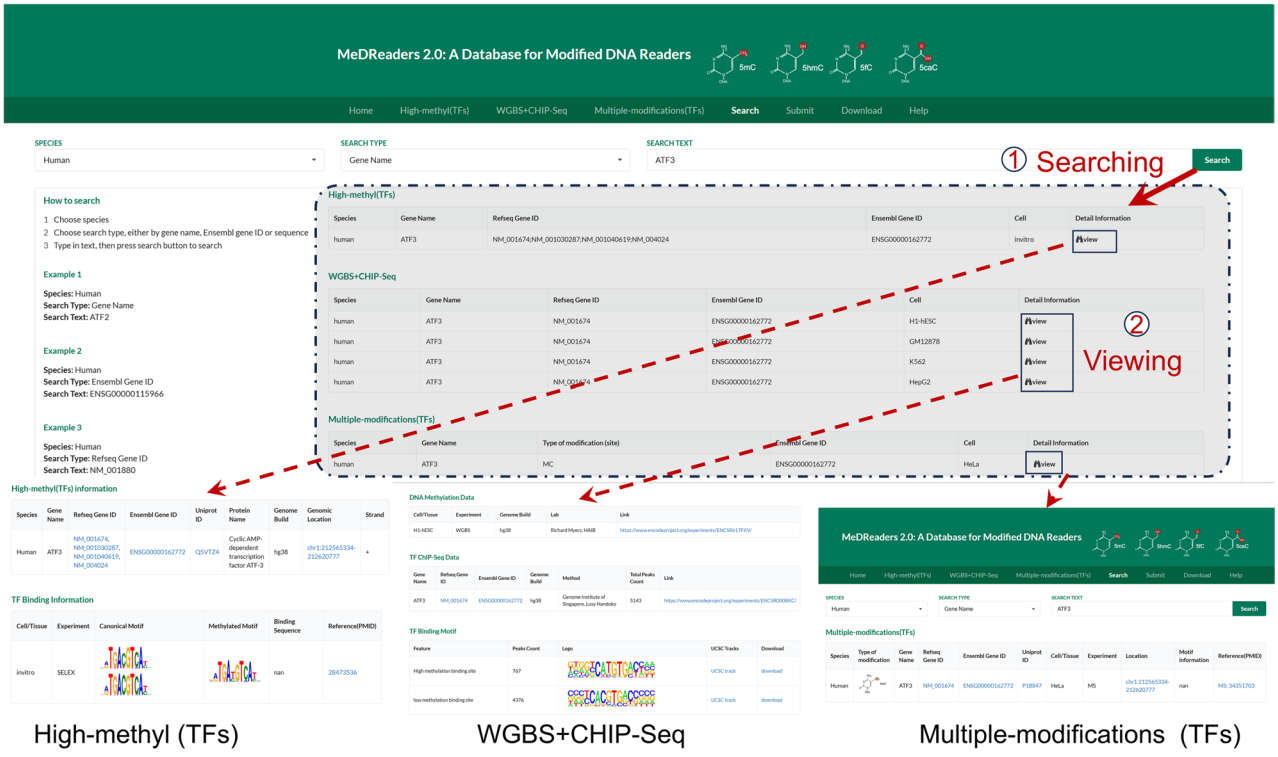

Fig. 2 Functionality of the search interface in MeDReaders 2.0

for further investigating the regulatory roles of specific cytosine modifications on TF binding during active DNA demethylation. Ultimately, understanding these diverse binding preferences will help shed light on the intricate interplay between dynamic epigenetic modifications and gene regulation.

Anticipating the rapid evolution of sequencing technologies and the continuous discovery of novel TF-DNA interactions, MeDReaders 2.0 provides a "Submit" page for users to upload new findings. All submitted data undergo rigorous validation based on supporting information, such as references (PMIDs) and experimental details. Interactions directly validated by experiments are promptly incorporated into the database. Meanwhile, raw data from TF ChIP-seq and cell line WGBS will first be processed through our established in silico pipeline prior to database integration. Through this continuous updating mechanism, we aim to provide a comprehensive resource that helps researchers further elucidate the mechanistic interplay between DNA modifications and TFs.

Browsing the database

MeDReaders 2.0 includes eight functional pages, three of which are dedicated to data presentation. The "High-5mC (TFs)" page summarizes TF information curated from the literature for high-5mC DNA. Users can explore TFs of interest by selecting a species and a gene name. The information provided includes Ensembl Gene ID, genomic location (coordinates), and motif information. The "WGBS+ChIP-seq" page displays predicted TF–5mC–DNA interactions derived from WGBS and ChIP-seq data. A notable feature of this page is the integration of UCSC Genome Browser tracks, enabling users to visualize the methylation status of TF-binding regions. In addition, the "Multiple-modifications (TFs)" page provides information on TF binding to cytosine modifications, including 5mC, 5hmC, and 5fC, annotated using MS. This page primarily presents TF identifier, modification types, and chromosomal positions.

Searching the database

To facilitate the examination of the conservation and cell-type specificity of TF binding to modified DNA, users can access the "Search" page to view comprehensive binding information for a specific TF. Figure 2 illustrates the binding of the human TF ATF3 to modified DNA across different cell lines. By clicking the "View" icon, users can access detailed information, which is consistent with the records on the "High-methyl (TFs)", "WGBS+ChIP-seq", and "Multiple-modifications (TFs)" pages.

Future development

The growing body of evidence suggests that epigenetic modifications play a crucial role in gene regulation by influencing transcription factor binding. However,

comprehensive and systematic studies that examine the detailed mechanisms underlying this regulation remain limited. The updated MeDReaders database continues to curate newly released datasets and published findings, thereby providing valuable insights. As additional data are integrated, MeDReaders will enable more precise identification of cell-specific effects of epigenetic modifications on TF binding. Moreover, incorporating additional epigenetic modifications into analyses of TF–DNA interactions will facilitate deeper investigation of TF regulatory mechanisms during active DNA methylation and demethylation dynamics.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-026-12852-2.

Supplementary Material 1.

Acknowledgements
Not applicable.

Authors' contributions
Conceptualization, H.F.L.; Methodology, H.F.L., M.R.Z.; Investigation, H.F.L., M.R.Z., and X.M.L.; Writing—original draft, H.F.L.; G.W.H and X.M.L.; Visualization, H.F.L. and X.M.L. All authors have read and agreed to the published version of the manuscript.

Funding
This work has been supported by the National Natural Science Foundation of China (62302342, 62371347) and the Zhejiang Provincial Postdoctoral Research Funding Project (ZJ2025176).

Data availability
All ChIP-seq and WGBS datasets analyzed in this study were obtained from the ENCODE and NCBI databases, and their corresponding accession numbers are documented in Supplementary Material 1. The results derived from these datasets can be accessed at https://bioinfor.nefu.edu.cn/MeDReaders2.0/.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors have read and approved the final manuscript and consent to its publication.

Competing interests
The authors declare no competing interests.

Received: 30 December 2025 / Accepted: 7 April 2026
Published online: 11 April 2026

References

1. Luo X, Zou Q. Identifying the stripe transcription factors and cooperative binding related to DNA methylation. Commun Biol. 2024;7(1):1265.
2. Li H, et al. Detection of transcription factors binding to methylated DNA by deep recurrent neural network. Brief Bioinform. 2022;23(1):bbab533.
3. Wang Y, et al. SBSM-Pro: support bio-sequence machine for proteins. Sci China-Information Sci. 2024;67(11):212106.
4. Kazemi E, et al. A genome-wide association study to identify candidate genes for erectile dysfunction. Brief Bioinform. 2021;22(4):bbaa338.
5. Liu X, et al. Cis-Regulatory Elements in Mammals. Int J Mol Sci. 2023;25(1):343.
6. Preissl S, Gaulton KJ, Ren B. Characterizing cis-regulatory elements using single-cell epigenomics. Nat Rev Genet. 2023;24(1):21–43.
7. Li Q, et al. RMNet: An RNA m6A Cross-species Methylation Detection Method for Nanopore Sequencing. Curr Drug Targets. 2025;26(11):799–812.
8. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003;33 Suppl:245–54.
9. Hendrich B, Tweedie S. The methyl-CpG binding domain and the evolving role of DNA methylation in animals. Trends Genet. 2003;19(5):269–77.
10. Yang Y, et al. Binding Analysis of Methyl-CpG Binding Domain of MeCP2 and Rett Syndrome Mutations. ACS Chem Biol. 2016;11(10):2706–15.
11. Liu Y, et al. Wemics: A Single-Base Resolution Methylation Quantification Method for Enhanced Prediction of Epigenetic Regulation. Adv Sci. 2024;11(21):230884.
12. Okano LM, et al. Identification of Epigenetic Regulatory Networks of Gene Methylation-miRNA-Transcription Factor Feed-Forward Loops in Basal-like Breast Cancer. Cells. 2025;14(16):1235.
13. Rimoldi M, et al. DNA methylation patterns of transcription factor binding regions characterize their functional and evolutionary contexts. Genome Biol. 2024;25(1):146.
14. Butz S, et al. DNA sequence and chromatin modifiers cooperate to confer epigenetic bistability at imprinting control regions. Nat Genet. 2022;54(11):1702–10.
15. Isbel L, Grand RS, Schübeler D. Generating specificity in genome regulation through transcription factor sensitivity to chromatin. Nat Rev Genet. 2022;23(12):728–40.
16. Hu S, et al. DNA methylation presents distinct binding sites for human transcription factors. Elife. 2013;2:e00726.
17. Prokhortchouk A, et al. The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev. 2001;15(13):1613–8.
18. Quenneville S, et al. In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol Cell. 2011;44(3):361–72.
19. Rishi V, et al. CpG methylation of half-CRE sequences creates C/EBPalpha binding sites that activate some tissue-specific genes. Proc Natl Acad Sci U S A. 2010;107(47):20311–6.
20. Balaramane D, et al. MethyLasso: a segmentation approach to analyze DNA methylation patterns and identify differentially methylated regions from whole-genome datasets. Nucleic Acids Res. 2024;52(21):e98.
21. Chen HY, Xu CH. Homeostasis of DNA Hemi-Methylation in Arabidopsis through Methylation Maintenance, DNA Replication, and Nucleosome Positioning Mechanisms. Advanced Science. 2025;12(38):e05808.
22. Huang HY, et al. Target enrichment of methylated circulating tumor DNA for colorectal cancer detection. Cancer Res. 2024;84(6):3656.
23. Ju X, et al. DNA methylation in poultry: a review. J Anim Sci Biotechnol. 2023;14(1):138.
24. Silva TC, et al. MethReg: estimating the regulatory potential of DNA methylation in gene transcription. Nucleic Acids Res. 2022;50(9):e51.
25. Li M, et al. Enriched transcription factor binding sites in hypermethylated gene promoters in drug resistant cancer cells. Bioinformatics. 2008;24(16):1745–8.
26. Yin Y, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science. 2017;356(6337):eaaj2239.
27. Brinkman AB, et al. Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. Genome Res. 2012;22(6):1128–38.
28. Song G, et al. An all-to-all approach to the identification of sequence-specific readers for epigenetic DNA modifications on cytosine. Nat Commun. 2021;12(1):795.
29. Wang G, et al. MeDReaders: a database for transcription factors that bind to methylated DNA. Nucleic Acids Res. 2018;46(D1):D146–51.
30. Strogantsev R, et al. Allele-specific binding of ZFP57 in the epigenetic regulation of imprinted and non-imprinted monoallelic expression. Genome Biol. 2015;16(1):112.
31. Mann IK, et al. CG methylated microarrays identify a novel methylated sequence bound by the CEBPB|ATF4 heterodimer that is active in vivo. Genome Res. 2013;23(6):988–97.
32. Spruijt CG, et al. Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. Cell. 2013;152(5):1146–59.
33. Bartke T, et al. Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell. 2010;143(3):470–84.

34. Bai L, et al. Proteome-Wide Profiling of Readers for DNA Modification. Adv Sci (Weinh). 2021;8(19):e2101426.
35. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27(11):1571–2.
36. Tsuji J, Weng Z. Evaluation of preprocessing, mapping and postprocessing algorithms for analyzing whole genome bisulfite sequencing data. Brief Bioinform. 2016;17(6):938–52.
37. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57–74.
38. Kagda MS, et al. Data navigation on the ENCODE portal. Nat Commun. 2025;16(1):9592.

Publisher's note