D1004–D1009 *Nucleic Acids Research, 2022, Vol. 50, Database issue* *Published online 30 October 2021*
https://doi.org/10.1093/nar/gkab972

EWAS Open Platform: integrated data, knowledge and toolkit for epigenome-wide association study

Zhuang Xiong[1,2,3,†], Fei Yang[1,2,3,†], Mengwei Li[1,2,†], Yingke Ma[1,2], Wei Zhao[1,2,3], Guoliang Wang[1,2,3], Zhaohua Li[1,2,3], Xinchang Zheng[1,2], Dong Zou[1,2], Wenting Zong[1,2,3], Hongen Kang[1,2,3], Yaokai Jia[1,2], Rujiao Li[1,2,*], Zhang Zhang[1,2,3,*] and Yiming Bao[1,2,3,*]

[1]National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences / China National Center for Bioinformation, Beijing 100101, China, [2]CAS Key Laboratory of Genome Sciences and Information, Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing 100101, China and [3]University of Chinese Academy of Sciences, Beijing 100049, China

Received September 14, 2021; Revised October 03, 2021; Editorial Decision October 04, 2021; Accepted October 20, 2021

ABSTRACT

Epigenome-Wide Association Study (EWAS) has become a standard strategy to discover DNA methylation variation of different phenotypes. Since 2018, we have developed EWAS Atlas and EWAS Data Hub to integrate a growing volume of EWAS knowledge and data, respectively. Here, we present EWAS Open Platform (https://ngdc.cncb.ac.cn/ewas) that includes EWAS Atlas, EWAS Data Hub and the newly developed EWAS Toolkit. In the current implementation, EWAS Open Platform integrates 617 018 high-quality EWAS associations from 910 publications, covering 51 phenotypes, 275 diseases and 104 environmental factors. It also provides well-normalized DNA methylation array data and the corresponding metadata from 115 852 samples, which involve 707 tissues, 218 cell lines and 528 diseases. Taking advantage of integrated knowledge and data in EWAS Atlas and EWAS Data Hub, EWAS Open Platform equips with EWAS Toolkit, a powerful one-stop site for EWAS enrichment, annotation, and knowledge network construction and visualization. Collectively, EWAS Open Platform provides open access to EWAS knowledge, data and toolkit and thus bears great utility for a broader range of relevant research.

INTRODUCTION

With the explosive growth of epigenome-wide association studies (EWAS), huge amounts of data and knowledge related to EWAS have been accumulated (1). Since these data hold great potential for clinical translations, a standardized platform for data archive, retrieval and exploration is indispensable. In order to identify potential biomarkers for human healthcare and disease treatment (2–6), a large number of EWAS associations have been reported in publications, posing great challenges in literature curation and knowledge synthesis. To this end, valuable efforts have been made worldwide, with the purpose to develop several databases and tools in aid of EWAS data integration and analysis (7–12). In 2018, we, the EWAS team of the National Genomics Data Center (NGDC) (13,14), launched EWAS Atlas (https://ngdc.cncb.ac.cn/ewas/atlas), a EWAS knowledgebase hosting manually curated high-quality EWAS associations (15). Subsequently, we further constructed EWAS Data Hub (https://ngdc.cncb.ac.cn/ewas/datahub), a data portal for collecting and normalizing DNA methylation array data as well as archiving associated metadata (16).

Since the first release of EWAS Atlas, there are over 28 000 visitors with a total of 127 000 accesses, along with >100 emails and phone calls from worldwide users (Table 1). To promote the data-to-bedside research to inform diagnosis and guide treatments by, for example, the identification of more indicative biomarkers from data (16), we developed EWAS Toolkit (https://ngdc.cncb.ac.cn/ewas/toolkit), a web-based tool suite for EWAS downstream analyses that comprise a series of online services for EWAS enrichment & annotation and network visualization, on the basis of integrated knowledge and data from EWAS Atlas, EWAS Data Hub, and the Roadmap Epigenomics Project (17). Pulling EWAS Atlas and EWAS Data Hub that have significant updates in the past several years, together with the recently developed EWAS Toolkit, here we introduce EWAS Open Platform (https://ngdc.cncb.ac.cn/ewas/) (Figure 1), serving as a one-stop site to deliver a portfolio of services for EWAS data, knowledge and toolkit.

[*]To whom correspondence should be addressed. Tel: +86 10 84097858; Fax: +86 10 84097720; Email: baoym@big.ac.cn
Correspondence may also be addressed to Rujiao Li. Email: lirj@big.ac.cn
Correspondence may also be addressed to Zhang Zhang. Email: zhangzhang@big.ac.cn
[†]The authors wish it to be known that, in their opinion, first three authors should be regarded as Joint First Authors.

Table 1. Main updates of EWAS open platform

	EWAS Atlas (2019)/EWAS Data Hub (2020)	EWAS Open Platform (2021) (EWAS Atlas/ EWAS Data Hub)
Data and Information		
Associations	329 172/NA	617 018/NA
Traits	305/NA	618/NA
Cohorts	1830/NA	3382/NA
Tissues or Cells	112/470	193/925
Studies	898/NA	1437/NA
Publications	649/NA	910/NA
Samples	NA/75 344	NA/115 852
Disease	128/306	275/528
Fields	NA/178	NA/242
Toolkit		
Enrichments	Trait, Genomic location, GO, KEGG and Motif enrichment	
Annotations	Chromatin state, Histone modification, Tissue methylation and Expression regulation	
Network	Knowledge graph	
Usage		
Visitors[a]	28 879	
Visits[a]	127 847	
Citations[b]	84	

[a] Data as of 12 September 2021.
[b] Data retrieved from Google Scholar, as of 12 September 2021.

MAJOR COMPONENTS AND UPDATES

EWAS Open Platform is an open platform for epigenome-wide association studies that incorporates three components: EWAS Data Hub for data collection and standardized normalization, EWAS Atlas for knowledge extraction and curation, and EWAS Toolkit for downstream analysis and visualization. Each component is a stand-alone database or web server.

EWAS Atlas

As a curated knowledgebase of EWAS Open Platform, EWAS Atlas has been enriched by adding a total of 287 864 EWAS associations manually curated from 509 publications, nearly doubling the number of associations by comparison with the first release in 2018. As a result, EWAS Atlas currently houses a total of 617 018 high-quality EWAS associations reported in 910 publications, involving 618 traits, 1437 studies, 3382 cohorts and 193 tissues/cell types (Table 1). Users can browse the associations in light of trait, probe, gene, study and publication, which are displayed in five panels on the browse page. In addition to efficient search engine and handy download services, EWAS Atlas provides expanded panels to show detailed and quantitative information, such as the rank, *P* value, effect size, trait and methylation value.

EWAS Data Hub

As a normalized data portal of EWAS Open Platform, EWAS Data Hub features comprehensive integration of all available datasets generated by Infinium HumanMethylation450 and MethylationEPIC BeadChip from GEO (18), TCGA (19), ArrayExpress (20) and ENCODE (21). In addition, it adopts a set of curation processes to eliminate batch effects and improve data quality (22). EWAS Data Hub has been significantly updated by including 40 508 high-quality samples of DNA methylation array data and metadata since the first release in 2019. Currently, a total of 115 852 samples are stored in EWAS Data Hub, covering 242 fields, 925 tissue/cells (including 218 cell lines) and 528 diseases (Table 1).

EWAS Toolkit

As an indispensable component of EWAS Open Platform, EWAS Toolkit is a new powerful one-stop analysis service for EWAS downstream analysis. Currently, EWAS Toolkit firstly features trait enrichment and network visualization by leveraging 617 018 high-quality associations from 910 publications in EWAS Atlas. It has been widely used in the retrieval and discovery of epigenetics biomarkers since its release (23–25). At the same time, combined with the methylation and expression profile data in the EWAS Data Hub, we provide tissue methylation and expression regulation annotations across 31 tissues. In addition, EWAS Toolkit integrates knowledge and data, organically combines EWAS Atlas and EWAS Data Hub, and provides users with a wide range of analysis and visualization including enrichment, annotation and network visualization (Figure 2).

Enrichment and annotation

EWAS Toolkit allows users to provide probes and trait term as input. Analyzed results are categorized into five enrichment sections (trait, genomic location, GO, KEGG and motif enrichment) and four annotation sections (chromatin state, histone modification, methylation and expression regulation annotation). DNA methylation probes related to vitamin B12 supplementation reported in the literature are provided as an example in the web page of EWAS Toolkit (26). To facilitate users for further analysis, all results can be packaged and downloaded.

Enrichment. Trait enrichment analysis is based on the curated associations related to traits in the EWAS Atlas. Weighted Fisher's exact test was used to compute the probability of co-occurrence between input DNA methylation probes and trait-related DNA methylation probes (Figure 2A). For genomic location enrichment, we pre-defined 13 location categories based on the location relative to gene and CpG island. The result of genomic location enrichment shows the enrichment of input sites in different regions of the genome (Figure 2B). Because the number of probes designed for each gene on the DNA methylation array is not equal, the use of traditional Fisher's exact test or chi-square test can lead to bias in GO and KEGG enrichment results (27,28). To address this issue, we adopted the algorithm designed by Phipson *et al*. based on the Wallenius non-central hypergeometric distribution (29), which can calculate and add the number of probes designed on each gene as prior knowledge to the inspection process. The 'gometh' function from the R package 'missMethyl' was used in this implementation (29). The GO and KEGG

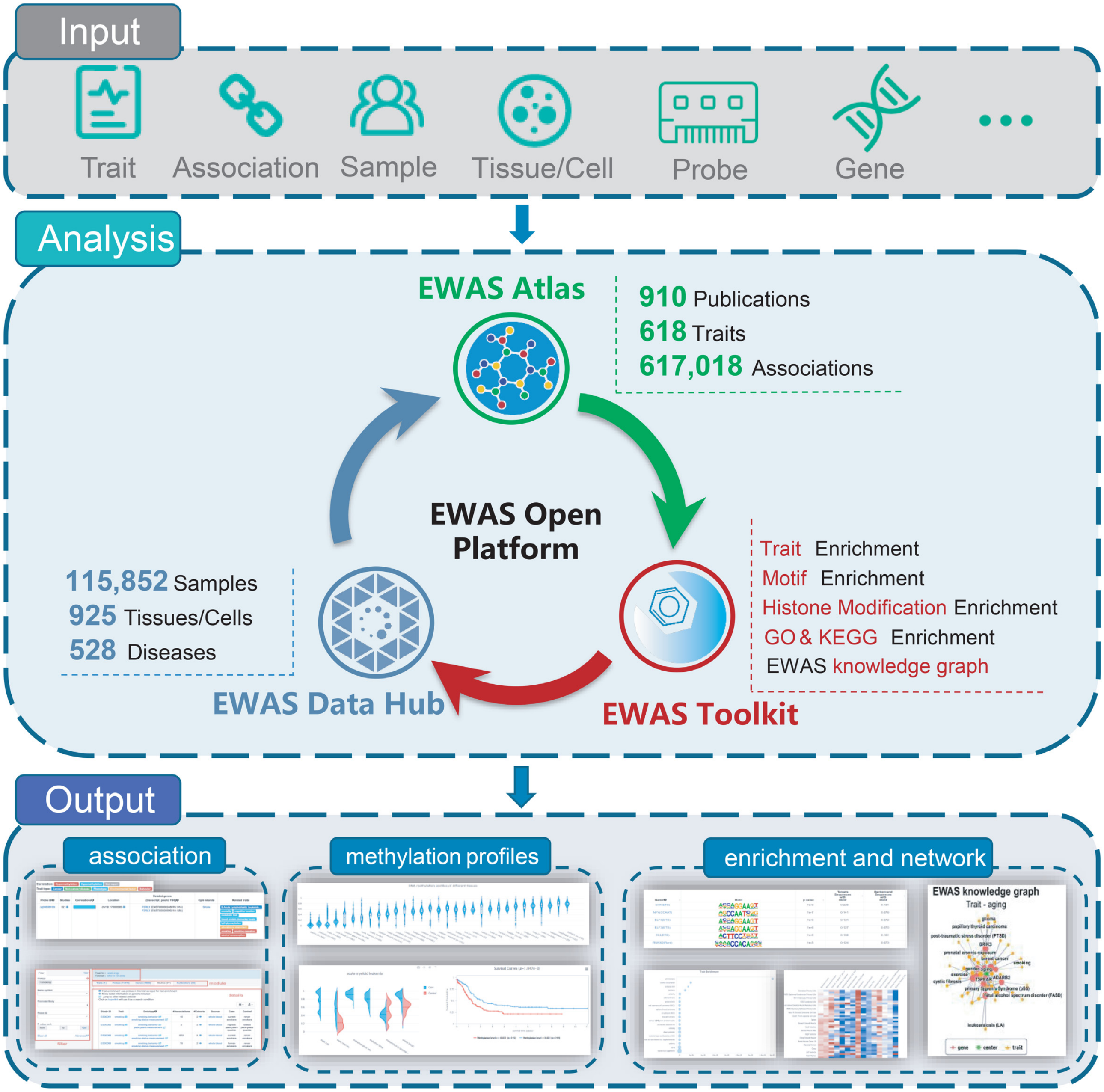
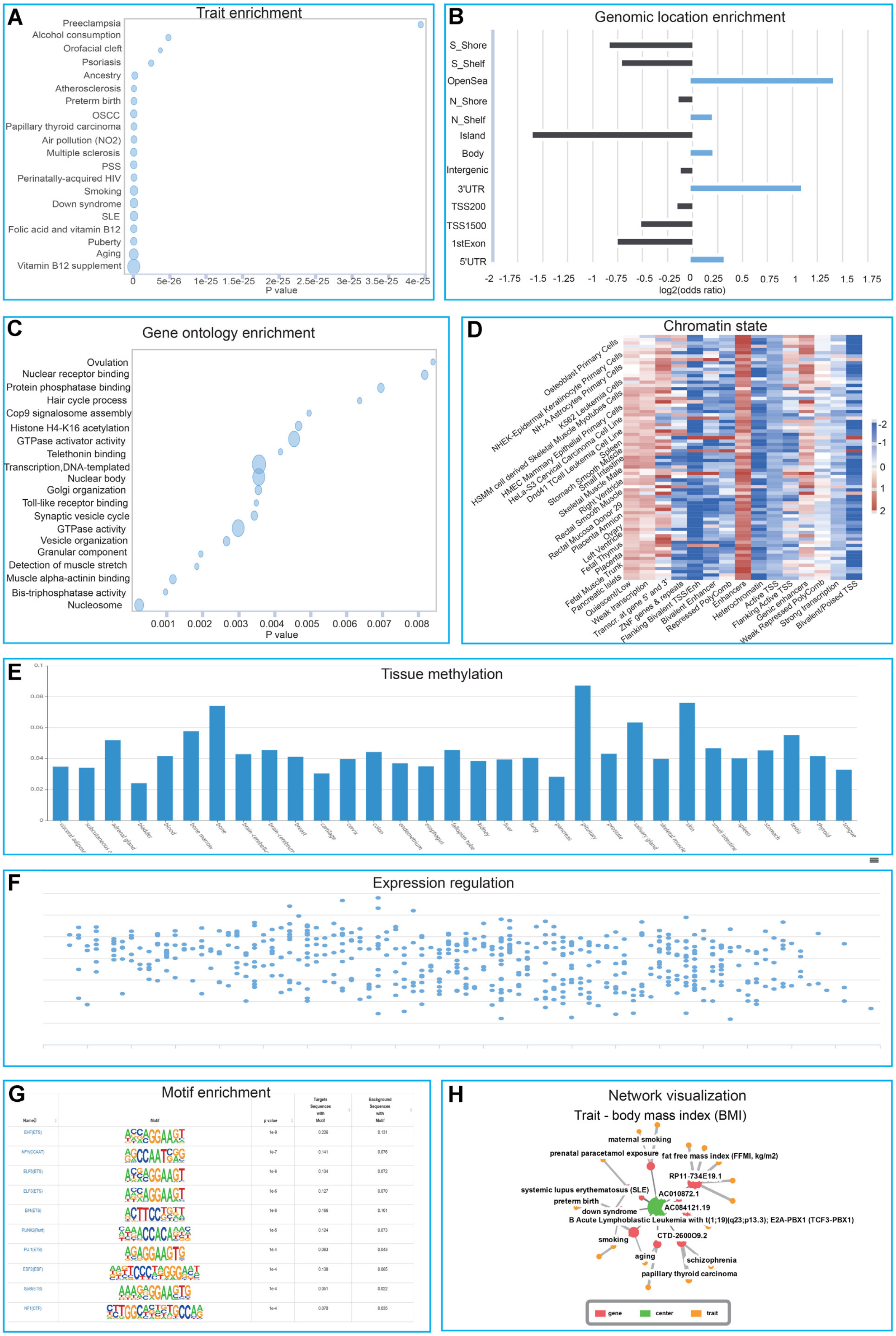


Figure 1. Schematic overview of EWAS Open Platform data processing workflow.

enrichment results show the enrichment of genes near the input probes in the GO entry and KEGG pathway, respectively (Figure 2C). Motif refers to a characteristic sequence with biological significance, such as binding sequence of regulatory factors like transcription factors. Motif enrichment near DNA methylation can aid the identification of proteins that interact with DNA methylation. The HOMER (Hypergeometric Optimization of Motif EnRichment, http://homer.ucsd.edu/homer) knowledgebase, which leverages a massive amount of chromatin immunoprecipitation data for transcription factor motif identification (30), was used for motif enrichment in EWAS Toolkit (Figure 2G).

Annotation. The state of histone modification, a chemical modification that occurs on histones, can influence the structure of chromatin, thereby affecting gene expression. For histone modification enrichment, we used the histone modification data of 127 cells and tissues in the Roadmap Epigenomics Project (17). Chromatin state refers to the regulatory function and current activity of a region of the genome. It is usually predicted by histone modification, DNA methylation, and gene expression. For the enrichment of chromatin states, we used the chromatin state data by the Roadmap Epigenomics Project to identify a total of 15 chromatin states (Figure 2D). EWAS Toolkit provides tissue methylation and expression regulation annotations based



Name	Motif	p value	Targets Sequences with Motif	Background Sequences with Motif
EHF(ETS)	ACCAGGAAGT	1e-9	0.226	0.131
NFY(CCAAT)	AGCCAATCGG	1e-7	0.141	0.076
ELF5(ETS)	ACGAGGAAGT	1e-6	0.134	0.072
ELF3(ETS)	ACCAGGAAGT	1e-6	0.127	0.070
Elf4(ETS)	ACTTCCTGTT	1e-6	0.166	0.101
RUNX2(Runt)	CAAACCACAAAC	1e-5	0.124	0.073
PU.1(ETS)	AGAGGAAGTG	1e-4	0.083	0.043
EBF2(EBF)	AAGTCCCTAGGGAAT	1e-4	0.138	0.085
SpiB(ETS)	AAAGAGGAAGTG	1e-4	0.051	0.022
NF1(CTF)	CTTGGCACTGTGCCAA	1e-4	0.070	0.035

Figure 2. Analysis panels of EWAS Toolkit. (**A**) Trait enrichment, showing that specific traits are significantly enriched. (**B**) Genomic location enrichment, demonstrating that DNA methylation sites related to vitamin B12 supplementation are enriched in Non-CpG island and 3'UTR regions. (**C**) Gene ontology enrichment, showing the enriched GO terms of genes near the input probe set. (**D**) Chromatin state, showing the significant enrichment in the enhancer region. (**E**) Tissue methylation, providing methylation levels across tissues of the input probe site. (**F**) Expression regulation, showing the relationship between the methylation level of the input site and the expression level of nearby genes. (**G**) Motif enrichment, displaying the enriched motifs and their associated transcription factors in the vicinity of the input probe set. (**H**) Network visualization, by taking BMI as an example.

on EWAS Data Hub. The annotation of tissue methylation displays the input probe's tissue-specificity (tau) (16) as well as the DNA methylation level in 31 tissues (Figure 2E). The relationship between methylation level of the input probes and expression level of nearby genes is depicted as expression regulation (Figure 2F).

Network visualization

In order to assist users to explore the hierarchical associations between traits and genes in a visualized network, EWAS Toolkit is capable to provide an EWAS knowledge graph by linking any given trait/gene to its associated genes/traits. Users can select one or two ways to calculate the correlation coefficient between genes and traits. The first is based on the number of publications reporting on the relations between a gene and a trait. The second is based on the inferred associations from data, taking the sum of the number of associations between all probes on the gene and traits as the association coefficient. On the knowledge graph page (https://ngdc.cncb.ac.cn/ewas/network), users can specify various parameters, such as the central node, the number of layers displayed, the maximum number of edges from each node, and the calculation method of the correlation coefficient. The network structure can be updated in real time when any parameter is changed (Figure 2H).

DISCUSSION AND FUTURE DEVELOPMENTS

In this study, we present EWAS Open Platform, an integrated open platform for EWAS data storage and download, knowledge collection and browsing, and downstream analysis and visualization. Apart from adding 287 864 EWAS associations from 509 publications to EWAS Atlas and 40 508 high-quality samples to EWAS Data Hub, EWAS Open Platform now contains a data analysis component, EWAS Toolkit, that supports various online analyses for EWAS enrichment, annotation, and network visualization. In particular, EWAS Toolkit has generated the first re-mining analysis platform based on enormous amounts of knowledge data from literatures, integrating trait enrichment and EWAS network visualization paired with knowledge graph. Future developments of EWAS Open Platform are frequent updates of these three components and enhancement of data sharing and information flow between EWAS Data Hub, EWAS Atlas and EWAS toolkit. For EWAS Atlas, we will optimize the curation model, accept community-curated annotations combined with expert review, and improve curation efficiency and quality. For EWAS data Hub, we will update the reference DNA methylation profile and improve cell line browsing and display based on the newly added cell line data. For EWAS Toolkit, we will use graph theory and machine learning methods to facilitate in-depth mining of knowledge graph, as well as analyze and predict complex relationships between phenotypes, environments, and behaviors. In addition, by combining methylation databases in NGDC, particularly MethBank (31) and scMethBank (32), EWAS Open Platform will provide easy and open access to more comprehensive data, knowledge and toolkit.

DATA AVAILABILITY

EWAS Open Platform is an integrated database and analysis platform for Epigenome-Wide Association Study which is freely available online and all data can be accessed at https://ngdc.cncb.ac.cn/ewas.

ACKNOWLEDGEMENTS

We thank a number of users for reporting bugs and providing suggestions.

FUNDING

Strategic Priority Research Program of the Chinese Academy of Sciences [XDB38030200, XDB38030400, XDA19050302]; Key Technology Talent Program of the Chinese Academy of Sciences; National Key Research and Development Program of China [2016YFE0206600, 2017YFC0908403, 2017YFC0907502]; Genomics Data Center Construction of Chinese Academy of Sciences [WX145XQ07-04]; The Professional Association of the Alliance of International Science Organizations [ANSO-PA-2020–07]; The Open Biodiversity and Health Big Data Programme of IUBS; National Natural Science Foundation of China [32030021, 31871328]. Funding for open access charge: Strategic Priority Research Program of the Chinese Academy of Sciences.

Conflict of interest statement. None declared.

REFERENCES

1. Wei,S., Tao,J., Xu,J., Chen,X., Wang,Z., Zhang,N., Zuo,L., Jia,Z., Chen,H., Sun,H. *et al.* (2021) Ten years of EWAS. *Adv Sci (Weinh)*, https://doi.org/10.1002/advs.202100727.
2. Flanagan,J.M. (2015) Epigenome-wide association studies (EWAS): past, present, and future. *Methods Mol. Biol.*, **1238**, 51–63.
3. Liang,X.Y., Justice,A.C., So-Armah,K., Krystal,J.H., Sinha,R. and Xu,K. (2021) DNA methylation signature on phosphatidylethanol, not on self-reported alcohol consumption, predicts hazardous alcohol consumption in two distinct populations. *Mol Psychiatr*, **26**, 2238–2253.
4. Schiele,M.A. and Domschke,K. (2018) Epigenetics at the crossroads between genes, environment and resilience in anxiety disorders. *Genes Brain Behav.*, **17**, e12423.
5. van Dongen,J., Zilhao,N.R., Sugden,K., Hannon,E.J., Mill,J., Caspi,A., Agnew-Blais,J., Arseneault,L., Corcoran,D.L., Moffitt,T.E. *et al.* (2019) Epigenome-wide association study of attention-deficit/hyperactivity disorder symptoms in adults. *Biol Psychiat*, **86**, 599–607.
6. Bragelmann,J., Ponce,C.B., Marcelain,K., Roessler,S., Goeppert,B., Gallegos,I., Colombo,A., Sanhueza,V., Morales,E., Rivera,M.T. *et al.* (2021) Epigenome-wide analysis of methylation changes in the sequence of gallstone disease, dysplasia, and gallbladder cancer. *Hepatology*, **73**, 2293–2310.
7. Liu,D., Zhao,L., Wang,Z., Zhou,X., Fan,X., Li,Y., Xu,J., Hu,S., Niu,M., Song,X. *et al.* (2019) EWASdb: epigenome-wide association study database. *Nucleic Acids Res.*, **47**, D989–D993.
8. Xu,J., Zhao,L., Liu,D., Hu,S., Song,X., Li,J., Lv,H., Duan,L., Zhang,M., Jiang,Q. *et al.* (2018) EWAS: epigenome-wide association study software 2.0. *Bioinformatics*, **34**, 2657–2658.
9. Xiong,Y., Wei,Y., Gu,Y., Zhang,S., Lyu,J., Zhang,B., Chen,C., Zhu,J., Wang,Y., Liu,H. *et al.* (2017) DiseaseMeth version 2.0: a major expansion and update of the human disease methylation database. *Nucleic Acids Res.*, **45**, D888–D895.
10. Ongenaert,M., Van Neste,L., De Meyer,T., Menschaert,G., Bekaert,S. and Van Criekinge,W. (2008) PubMeth: a cancer methylation database combining text-mining and expert annotation. *Nucleic Acids Res.*, **36**, D842–D846.

11. Breeze,C.E., Reynolds,A.P., van Dongen,J., Dunham,I., Lazar,J., Neph,S., Vierstra,J., Bourque,G., Teschendorff,A.E., Stamatoyannopoulos,J.A. *et al.* (2019) eFORGE v2.0: updated analysis of cell type-specific signal in epigenomic data. *Bioinformatics*, **35**, 4767–4769.
12. Huang,H.Y., Li,J., Tang,Y., Huang,Y.X., Chen,Y.G., Xie,Y.Y., Zhou,Z.Y., Chen,X.Y., Ding,S.Y., Luo,M.F. *et al.* (2021) MethHC 2.0: information repository of DNA methylation and gene expression in human cancer. *Nucleic Acids Res.*, **49**, D1268–D1275.
13. CNCB-NGDC Members and Partners (2021) Database Resourcesof the National Genomics Data Center, China National Center for Bioinformationin 2022. *Nucleic Acids Res.*, https://doi.org/10.1093/nar/gkab951.
14. CNCB-NGDC Members and Partners (2021) Database resources of the National Genomics Data Center, China National Center for Bioinformation in 2021. *Nucleic Acids Res.*, **49**, D18–D28.
15. Li,M., Zou,D., Li,Z., Gao,R., Sang,J., Zhang,Y., Li,R., Xia,L., Zhang,T., Niu,G. *et al.* (2019) EWAS Atlas: a curated knowledgebase of epigenome-wide association studies. *Nucleic Acids Res.*, **47**, D983–D988.
16. Xiong,Z., Li,M., Yang,F., Ma,Y., Sang,J., Li,R., Li,Z., Zhang,Z. and Bao,Y. (2020) EWAS Data Hub: a resource of DNA methylation array data and metadata. *Nucleic Acids Res.*, **48**, D890–D895.
17. Roadmap Epigenomics,C., Kundaje,A., Meuleman,W., Ernst,J., Bilenky,M., Yen,A., Heravi-Moussavi,A., Kheradpour,P., Zhang,Z., Wang,J. *et al.* (2015) Integrative analysis of 111 reference human epigenomes. *Nature*, **518**, 317–330.
18. Barrett,T., Wilhite,S.E., Ledoux,P., Evangelista,C., Kim,I.F., Tomashevsky,M., Marshall,K.A., Phillippy,K.H., Sherman,P.M., Holko,M. *et al.* (2012) NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.*, **41**, D991–D995.
19. Hutter,C. and Zenklusen,J.C. (2018) The Cancer Genome Atlas: creating lasting value beyond its data. *Cell*, **173**, 283–285.
20. Parkinson,H., Sarkans,U., Shojatalab,M., Abeygunawardena,N., Contrino,S., Coulson,R., Farne,A., Lara,G.G., Holloway,E., Kapushesky,M. *et al.* (2005) ArrayExpress - a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.*, **33**, D553–D555.
21. Davis,C.A., Hitz,B.C., Sloan,C.A., Chan,E.T., Davidson,J.M., Gabdank,I., Hilton,J.A., Jain,K., Baymuradov,U.K., Narayanan,A.K. *et al.* (2018) The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.*, **46**, D794–D801.
22. Xiong,Z., Li,M., Ma,Y., Li,R. and Bao,Y. (2021) GMQN: A reference-based method for correcting batch effects as well as probes bias in HumanMethylation BeadChip. bioRxiv doi: http://doi.org/10.1101/2021.09.06.459116, 07 September 2021, preprint: not peer reviewed.
23. Zhu,T., Gao,Y., Wang,J., Li,X., Shang,S., Wang,Y., Guo,S., Zhou,H., Liu,H., Sun,D. *et al.* (2019) CancerClock: a DNA methylation age predictor to identify and characterize aging clock in Pan-cancer. *Front. Bioeng. Biotechnol.*, **7**, 388.
24. Jimenez,B.P., Liu,F., Caliebe,A., Gonzalez,D.M., Bell,J.T., Kayser,M. and Vidaki,A. (2021) Equivalent DNA methylation variation between monozygotic co-twins and unrelated individuals reveals universal epigenetic inter-individual dissimilarity. *Genome Biol.*, **22**, 18.
25. Gatev,E., Inkster,A.M., Negri,G.L., Konwar,C., Lussier,A.A., Skakkebaek,A., Sokolowski,M.B., Gravholt,C.H., Dunn,E.C., Kobor,M.S. *et al.* (2021) Autosomal sex-associated co-methylated regions predict biological sex from DNA methylation. *Nucleic Acids Res.*, **49**, 9097–9116.
26. Yadav,D.K., Shrestha,S., Lillycrop,K.A., Joglekar,C.V., Pan,H., Holbrook,J.D., Fall,C.H.D., Yajnik,C.S. and Chandak,G.R. (2018) Vitamin B-12 supplementation influences methylation of genes associated with Type 2 diabetes and its intermediate traits. *Epigenomics-UK*, **10**, 71–90.
27. Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T. *et al.* (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
28. Kanehisa,M. and Goto,S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.*, **28**, 27–30.
29. Phipson,B., Maksimovic,J. and Oshlack,A. (2016) missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics*, **32**, 286–288.
30. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
31. Li,R., Liang,F., Li,M., Zou,D., Sun,S., Zhao,Y., Zhao,W., Bao,Y., Xiao,J. and Zhang,Z. (2018) MethBank 3.0: a database of DNA methylomes across a variety of species. *Nucleic Acids Res.*, **46**, D288–D295.
32. Zong,W., Kang,H., Xiong,Z., Ma,Y., Jin,T., Gong,Z., Yi,L., Zhang,M., Wu,S., Wang,G. *et al.* (2021) scMethBank: a database for single-cell whole genome DNA methylation maps. *Nucleic Acids Res.*, https://doi.org/10.1093/nar/gkab833.