Grass Research

RESEARCH HIGHLIGHT

Open Access

https://doi.org/10.48130/grares-0026-0013

Grass Research **2026**, 6: e014

scPlantReg for single-cell epigenomics in complex forage grasses

Song Li[1*] and Mao Xia[2]

[1] *School of Plant and Environmental Sciences, Virginia Tech, Virginia 24060, USA*
[2] *College of Grassland Science and Technology, Sichuan Agricultural University, Chengdu, Sichuan 611130, China*
* Correspondence: songli@vt.edu (Li S)

Single-cell chromatin accessibility sequencing (scATAC-seq) has greatly advanced the study of gene regulatory landscapes, supporting high-resolution analysis of chromatin states and *cis*-regulatory element activity at the cell-type level. Forage grasses are important global crops with large and complex genomes, high heterozygosity, and incomplete genomic resources. These characteristics make single-cell epigenomic analysis especially challenging. Most available scATAC-seq tools are developed for mammalian systems. These tools rely on assumptions that do not match the features of plant genomes, and often require extensive manual adjustment[1]. To connect chromatin accessibility and real biological function, specialized computational tools and unified reference resources are needed for many plants[2]. These computational challenges hinder progress in forage grass research and limit the use of single-cell data in breeding for higher yield and stronger stress resistance. To fill this gap, Yan et al. have developed a reliable and integrated analytical framework called scPlantReg (Fig. 1) for research into single-cell epigenetic regulation in plants, particularly for a model forage grass, pearl millet (*Pennisetum glaucum*)[3].

The scPlantReg website provides a complete end-to-end solution for plant-related single-cell chromatin accessibility analysis. The framework integrates many key analytical steps into a consistent and reproducible pipeline, supported by the core Socrates2 module. The Socrates2 module includes tools for raw data processing, quality control, cell clustering, cell-type annotation, transcription factor motif enrichment, functional prediction of regulatory regions, and cross-species comparison. This integrated design is especially suitable for plant-specific scATAC-seq data, which often have high noise levels and are sensitive to the processing method[4]. Using standardized procedures and plant-specific parameters, Socrates2 helps reduce technical variation, improves the result stability, and lowers the barrier for high-quality scATAC-seq analysis in nonmodel plants. It performs well in forage grasses such as pearl millet, even with complex genomes and limited annotations. Its successful application in pearl millet demonstrates the reliability of the method and supports the broad feasibility of scPlantReg for the analysis of other grass species.

Another important technical component of scPlantReg is ScATACtor, a tool for cell-type annotation of plants' scATAC-seq data. A key difficulty in single-cell data annotation in plants is that many tissues contain closely related cell types that are hard to distinguish, partly because of the limited marker genes available for nonmodel species. ScATACtor uses supervised machine learning to identify cell types on the basis of genome-wide chromatin accessibility patterns. The method uses a random forest classifier trained on high-quality reference data and does not depend on a small set of marker regions. ScATACtor shows better performance than several existing annotation tools and achieves high accuracy and stable performance across different datasets. In addition, this method also supports cross-modality annotation between scATAC-seq and scRNA-seq data. The combined use of automated prediction and marker-based verification provides reliable cell-type labels, supporting consistent annotation across tissues and treatments for nonmodel grass species.

The authors tested scPlantReg using pearl millet, an important forage and grain grass with strong heat tolerance. Pearl millet has a complex genome, high heterozygosity, and limited public genomic resources compared with major crops. The team applied scPlantReg and successfully built the first single-cell chromatin accessibility landscape for pearl millet, including multiple developmental stages of leaves and roots. This analysis identified distinct cell types and more than 70,000 accessible chromatin regions, with many regions showing cell-type-specific patterns. The framework further supported functional discovery. For example, WRKY transcription factors were identified as potential regulators of xylem development. This function of WRKY has not been extensively reported in grasses. The authors focused on WRKY22 and predicted its target regulatory regions in xylem cells. These regions are related to genes involved in lignin biosynthesis and cell wall formation. Furthermore, these regulatory roles were confirmed by experimental assays. These results show that scPlantReg can turn single-cell epigenomic data into meaningful biological insights. It supports the discovery of key genes and regulatory elements for improving forage grasses. The authors also used scPlantReg to study chromatin dynamics under heat stress in pearl millet. They generated scATAC-seq data for control and heat-treated seedlings with two biological replicates, followed by strict quality filtering and batch correction to retain only high-quality cells for downstream analysis. The researchers analyzed changes in cell-type composition and found that vascular cell types showed moderate enrichment under heat stress. With consistent quality metrics and stable cell type annotations between replicates, these results suggest that observed changes are likely caused by real biological responses rather than technical bias. The framework also identified stress-specific accessible chromatin regions. Proliferating cells showed the highest number of treatment-specific regions, and more regions gained accessibility than lost accessibility under heat stress. This analysis demonstrates that scPlantReg can reliably resolve cell-type-specific regulatory responses to environmental stress and provides a practical model for studying abiotic stress adaptations in forage grasses.

One of the most important contributions of this work to the plant-related single-cell community is that the scPlantReg website includes a unified database of standardized plant scATAC-seq data. The database integrates reprocessed data from many tissues and developmental stages of important grasses such as rice (*Oryza sativa*), maize (*Zea mays*), sorghum (*Sorghum bicolor*), and pearl millet. Public single-cell epigenomic data from plants are often scattered and processed inconsistently. These issues prevent effective reuse and comparison. scPlantReg solves this problem by processing all data with the same pipeline. It unifies cell-type annotation and integrates additional epigenetic information. The

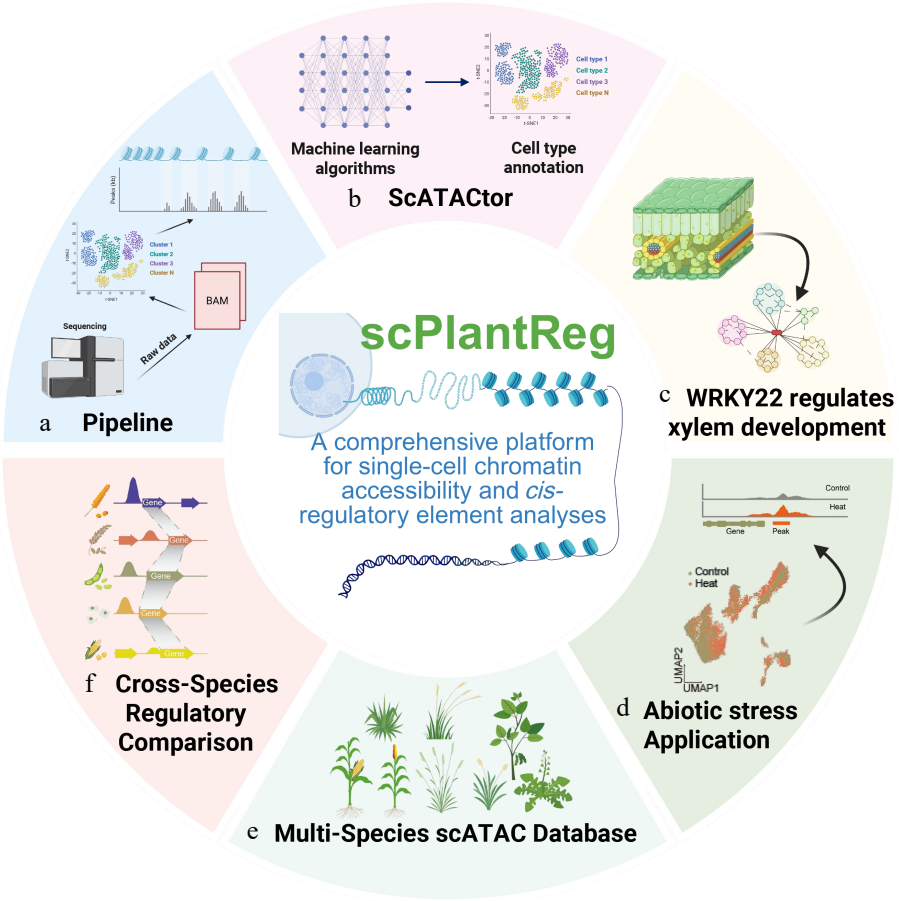

Fig. 1 Six key features of applications within the scPlantReg platform. (a) scPlantReg provides an end-to-end framework for plant-related single-cell chromatin accessibility analysis, integrating data processing, quality control, clustering, cell-type annotation, motif enrichment, regulatory prediction, and cross-species comparison through the Socrates2 module. The platform is optimized for scATAC-seq data from plants and improves reproducibility and stability, particularly in nonmodel species such as pearl millet. (b) ScATACtor, a machine learning-based tool for accurate cell-type annotation using genome-wide accessibility patterns, with support for cross-modality integration. (c) Application in pearl millet enabled the construction of a single-cell chromatin accessibility landscape, the identification of cell-type-specific regulatory regions, and functional validation of transcription factors such as WRKY22 in xylem development. (d) The framework also supports the analysis of stress responses, revealing cell-type-specific chromatin dynamics under heat stress. (e) scPlantReg hosts a standardized multispecies scATAC-seq database, integrating datasets from major grasses, thus enabling comparative analysis. (f) Cross-species analyses highlight conserved regulatory programs at the level of chromatin accessibility and transcription factor motifs, even in the absence of strong sequence conservation. Overall, the platform links single-cell epigenomic data to functional regulatory insights and supports gene discovery and molecular breeding in forage grasses. Created with BioRender (https://BioRender.com/a16u6sj).

website enables visual exploration, gene accessibility analysis, dynamic trajectory profiling, transcription factor motif enrichment, and cross-species comparison, supporting large-scale comparative regulatory analyses in plants and grasses. To demonstrate the utility of the website, the authors used the database to compare regulatory patterns across grass species in order to identify conserved regulatory programs during evolution. The results showed that regulatory conservation is often more evident in functional patterns than in exact DNA sequence similarity. Cell-type-specific accessibility and associated transcription factor motifs are preserved across species even when sequence conservation is low. For example, some shared accessible regions are enriched for AP2/EREBP family motifs, and these regions are linked to cell wall formation pathways that are conserved across grasses. Cell-type-specific shared regions show consistent enrichment of motifs that match the corresponding cell identities. These findings demonstrate that single-cell epigenomics can reveal functional conservation missed by traditional methods, offering a new way to study core developmental programs in forage grasses.

scPlantReg represents an important advance in the integrative and functional analysis of scATAC-seq data from plants. It unites multiple analytical layers into a single system optimized for plant genomes and performs reliably in nonmodel species with complex genomes. This work demonstrated robust analysis of developmental and stress-responsive data in pearl millet, a key forage grass. The large dataset and computational analysis enable the discovery and experimental validation of key regulatory factors. For the single-cell community, scPlantReg provides a unified database for cross-species research, and turns single-cell chromatin accessibility data into cell-type-specific regulatory insights that support crop improvement. In the future, this website will facilitate the identification of candidate regulatory elements that can be used in precision genome editing to improve yield, quality, and stress resilience. The platform is a starting point to support the continued growth of single-cell research in forage grasses and opens new opportunities for functional regulatory genomics and molecular breeding.

Author contributions

The authors confirm their contributions to the paper as follows: writing the manuscript: Li S; drawing figure: Xia M. Both authors reviewed the results and approved the final version of the manuscript.

Data availability

Data sharing is not applicable to this article, as no datasets were generated or analyzed during the current study.

Acknowledgments

The authors thank Dr. Linkai Huang from Sichuan Agricultural University for his assistance in writing this manuscript. Li S is supported by the Plant Genome Research Program (PGRP, Grant No. 2344169).

Conflict of interest

The authors declare that they have no conflict of interest.

Dates

Received 2 May 2026; Revised 2 May 2026; Accepted 19 May 2026; Published online 31 May 2026

References

[1] Stuart T, Srivastava A, Madad S, Lareau CA, Satija R. 2021. Single-cell chromatin state analysis with Signac. *Nature Methods* 18:1333–1341

[2] Preissl S, Gaulton KJ, Ren B. 2023. Characterizing *cis*-regulatory elements using single-cell epigenomics. *Nature Reviews Genetics* 24:21–43

[3] Yan H, Jin Y, Wang C, Zhang X, Jia J, et al. 2026. Single-cell chromatin accessibility and *cis*-regulatory element analyses in plants using the scPlantReg platform. *Nature Plants* 12:1116–1132

[4] Dorrity MW, Alexandre CM, Hamm MO, Vigil AL, Fields S, et al. 2021. The regulatory landscape of *Arabidopsis thaliana* roots at single-cell resolution. *Nature Communications* 12:3334