scientific data

Check for updates

OPEN

DATA DESCRIPTOR

Single-cell transcriptomic datasets of the amphibious plant Water wisteria

Gaojie Li[1,2], Abeer Kazmi[1,2] & Hongwei Hou[1,2] ✉

Global climate changes cause severe floods that hinder plant development and breeding. The phenomenon in which a single plant can significantly alter its leaf shape structure and physiology based on surrounding environments is known as heterophylly, a trait commonly observed in amphibious plants. Heterophylly is a special survival strategy that enables adaptation to both submerged and terrestrial conditions, aiding in response to environmental changes. Molecular mechanisms of heterophylly are not well understood due to a lack of models. *Hygrophila difformis*, also known as Water wisteria, shows typical heterophylly in terrestrial and submerged environments, and is an emerging model plant to explore heterophylly. In this study, we conducted single-cell RNA sequencing (scRNA-seq) analysis of isolated protoplasts from both terrestrial and submerged shoots of *H. difformis* to obtain data on the gene expression patterns of tissue-specific cells. These scRNA-seq data enabled the identification of gene expression patterns by cell types, providing a foundation to elucidate the molecular mechanisms of heterophylly and plant environmental adaptations at a single-cell level.

Background & Summary

In recent years, plants have been facing severe risks of drought and flooding caused by global climate change. Many terrestrial model plants, like Arabidopsis, tomato and maize, cannot afford long-term submergence, limiting their applications in this area. Amphibious plants, which grow in natural wetlands and can survive in both aquatic and aerial environments, are ideal models for studying plant adaptations to terrestrial and submerged conditions[1]. Heterophylly is the phenomenon where plants develop noticeably different leaf shapes depending on their surrounding environments. This occurrence is commonly observed in amphibious plants. To adapt to water level changes, many model plants develop adaptive strategies, including the low-oxygen quiescence strategy (LOQS), which conserves metabolism and growth during flooding, and the low-oxygen escape strategy (LOES), where plants quickly elongate their stem or petiole to reach the water surface. Recently, heterophylly was termed the third strategy to survive underwater. To adapt to the fluctuating water level, heterophyllous species develop complex, pinnate, or narrow leaves in submerged conditions, while they develop simple, serrated or broad leaves in terrestrial conditions[2–4]. Many environmental factors, including emergence/submergence[5], light quality[6,7], light intensity[8], and temperature[9] affect the heterophylly of various amphibious species. Phytohormones, such as abscisic acid (ABA), gibberellin (GA), cytokinin (CK), and ethylene, are also involved in the heterophylly regulation[10–14]. Recent studies using transcriptomic analysis have shown changes in gene expression under various conditions in some amphibious plants[2,15–19]. However, molecular mechanisms of heterophylly are still largely unknown. It was verified that the final leaf form depends on the early stage of shoots, especially the cell differentiation stage of leaf primordia, which initiates at the flanks of the shoot apical meristem (SAM)[8,20]. Therefore, it is important and timely to reveal the molecular changes of amphibious plants' shoots at a single-cell level.

Hygrophila difformis, commonly known as Water wisteria, is an emerging model plant for studies on heterophylly and plant adaptations to amphibious environments[20]. *H. difformis* develops simple and serrated leaves in terrestrial conditions, while it has complex and pinnate leaves underwater. Easy vegetative propagation and successfully established genetic transformation systems make it an ideal model plant to explore heterophylly at the genomic level[11,20]. Recent physiological studies on *H. difformis* also revealed advancements in photosynthetic

[1]The State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei, 430072, China. [2]University of Chinese Academy of Sciences, Beijing, 100049, China. ✉e-mail: houhw@ihb.ac.cn

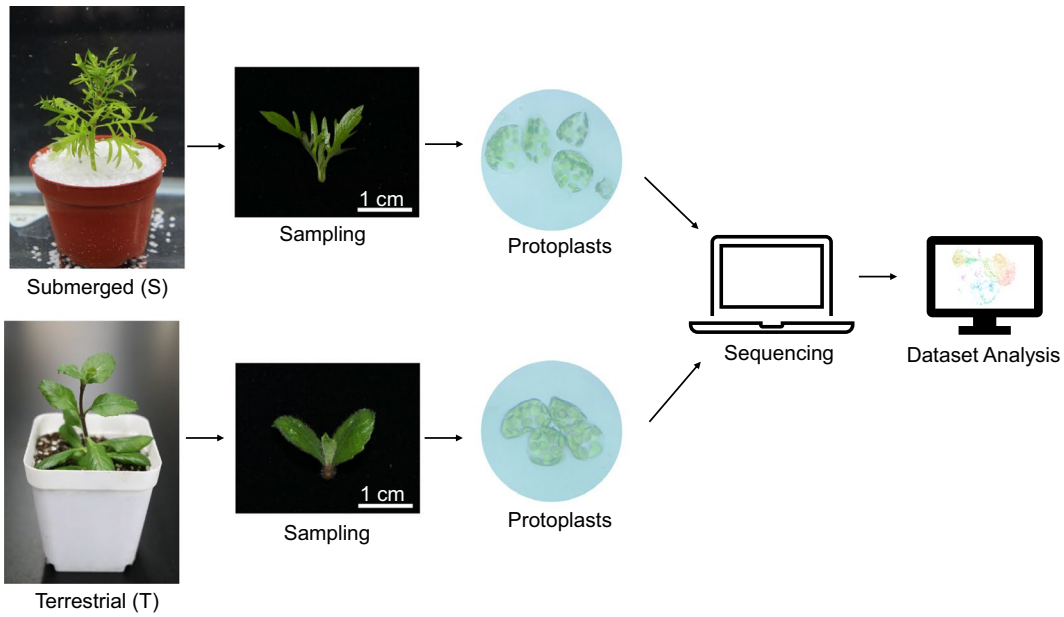

Fig. 1 Schematic overview of experimental design.

plasticity and its role as heavy metal indicators[21,22]. Additionally, the ethanol extract of *H. differmis* showed significant analgesic properties[23], indicating its bioactive potential.

Single-cell RNA sequencing (scRNA-seq) enables accurate sequencing and analysis of transcriptomes from individual plant cells, facilitating precise cell classification and exploration of their physiological functions and developmental characteristics[24–26]. Advancements in sequencing technologies have made scRNA-seq a powerful tool in developmental biology, applied to various terrestrial plant species, including Arabidopsis, soybean, tomato, and rice[27–29]. Currently, there is a lack of single-cell RNA sequencing (scRNA-seq) data for aquatic and amphibious plants.

In this study, we report the first scRNA-seq database of the amphibious plant *H. differmis* under terrestrial and submerged conditions from shoot protoplasts, providing valuable resources for future studies on heterophylly at a single-cell level. We conducted marker gene identification, gene cluster analysis, and expression analysis, along with clustering differentially expressed genes (DEGs) between terrestrial and submerged conditions. These datasets offer opportunities to understand the mechanisms of heterophylly in amphibious species at a single-cell level. Our study also reveals expression patterns of various genes in different cell type clusters, which could be strong candidates for future crop engineering in response to climate change.

Materials and Methods

Plant material, protoplast isolation, library construction, and scRNA-seq. *H. differmis* plants were grown in a growth chamber under white light conditions (23 °C, 16 h of light/8 h of dark, 30% relative humidity, light intensity of 60 μmol m^{-2} s^{-1}). Terrestrial plants planted in a soil mixture composed of peat soil, perlite and vermiculite (1:1:1), and irrigated every 7 days. Submerged plants were grown in aquariums (35 × 25 × 35 cm), filled with 30 cm depth tap water. 1/100 Murashige and Skoog medium (MS medium) was added in both conditions every two weeks. Plants kept growing in each condition for at least 2 months before sampling.

Protoplasts were isolated from apical shoots (about 1 cm in length), which were harvested from three individual plants that developed fully complex or serrated leaves under submerged or terrestrial environments (Fig. 1). A defined number of shoots were pooled per sample. Subsequently, samples were cut into 1 mm pieces and then immersed in protoplast isolation buffer [2% (w/v) cellulose R-10 (Yakult, Japan) and 0.5% (w/v) macerozyme (Yakult, Japan), 20 mM MES, 10 mM $CaCl_2$, 0.4 M mannitol, pH 5.8]. The harvested shoots were placed in a 50 mL syringe with 20 mL enzyme buffer, shaking at 100 rpm for 3 hours at room temperature. Samples were then filtered by a 200-nylon mesh filter, and filtered protoplasts were centrifuged at 1000 rpm for 3 minutes and washed with washing buffer [20 mM MES, 10 mM $CaCl_2$, 0.6 M mannitol, pH 5.8]. The protoplasts were stained with 0.2% trypan blue to check concentration and viability with a hemocytometer under a light microscope, and the live protoplasts were immediately loaded into the 10X Genomics Chromium Single Cell 3′ Platform (10X Genomics, USA). Subsequently, Single-cell suspensions were loaded onto prepared microfluidic chips to generate Gel Bead-In-Emulsions (GEMs), where cellular mRNA hybridized with barcoded oligonucleotides on gel beads. Reverse transcription generated cDNA with unique cell barcodes (CellBarcode) and unique molecular identifiers (UMI). Finally, libraries were constructed following the Chromium Single Cell 3′ Reagent Kits v3.1 protocol and sequenced on an Illumina NovaSeq. 6000 platform (150-bp paired-end mode). Single cell library preparation, sequencing, and primary data processing were conducted by Shanghai Origingene Bio-pharm Technology Co., Ltd (Shanghai, China).

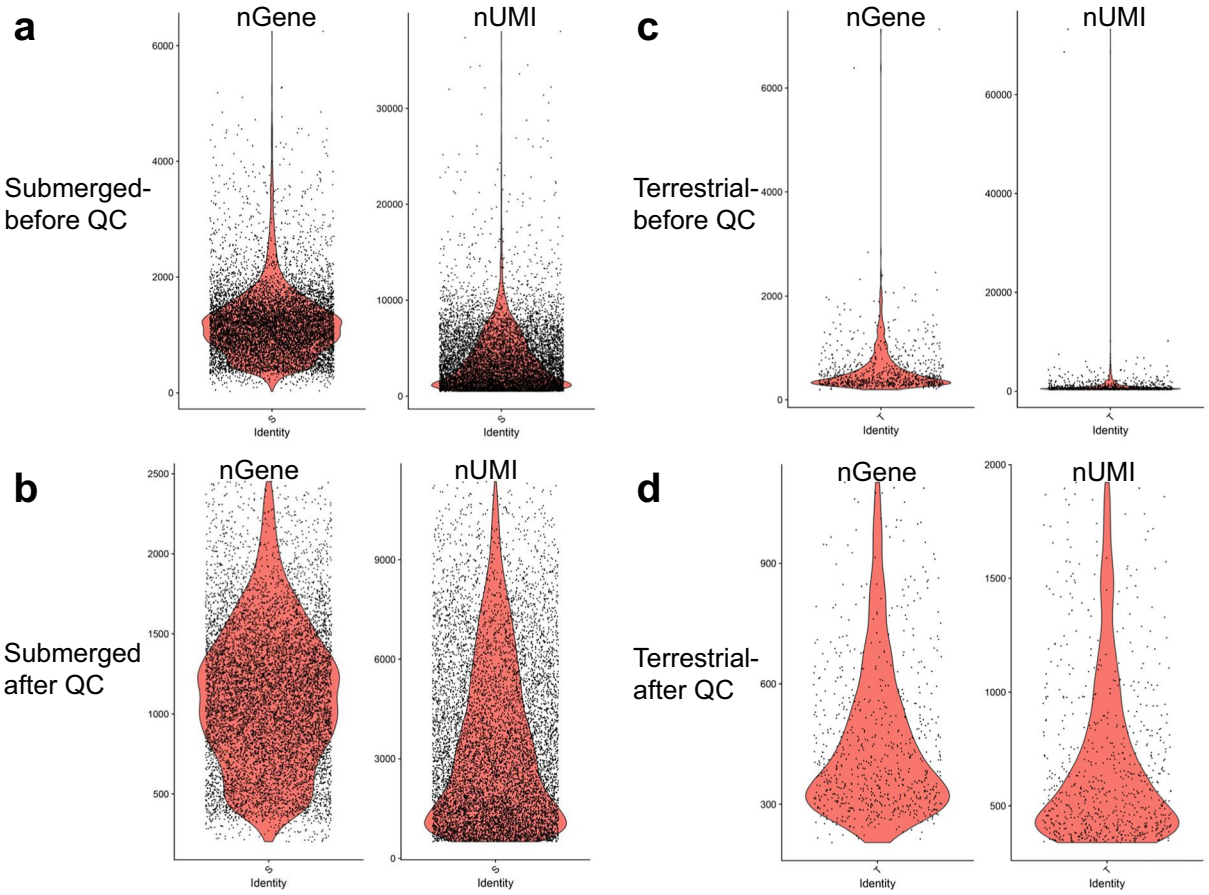

Fig. 2 Violin plots representing the nGene and nUMI in the raw and filtered scRNA-seq data. (**a**) Violin plots of the nGene and nUMI in the raw scRNA-seq data of submerged samples. (**b**) Violin plots of the nGene and nUMI in the filtered scRNA-seq data of submerged samples. (**c**) Violin plots of the nGene and nUMI in the raw scRNA-seq data of terrestrial samples. (**d**) Violin plots of the nGene and nUMI in the filtered scRNA-seq data of terrestrial samples.

Data processing and analysis. Raw fastaq files generated by 10X Genomics were aligned and quantified using FastQC (v0.11.7), and Cell Ranger (v6.1.1) was utilized to align reads to the *Hygrophila difformis* genome[30] using STAR (v2.7.1a). To identify homologous genes between *H. difformis* and *Arabidopsis thaliana*, proteins of *H. difformis* and *A. thaliana* were clustered using OrthoFinder (v.2.5.4)[31]. UMI counts per gene-cell pair were compiled into a gene expression matrix. For cell filtering, low-quality cells were excluded using Seurat (v4.0.3) with thresholds: Gene counts: Mean $\pm$ 2 SD; Mitochondrial genes: $<$30%; Ribosomal genes: $<$30%. UMI counts were log-normalized, and PCA analysis was performed on highly variable genes (top 2,000). Shared Nearest Neighbor (SNN) clustering at resolution 0.8 was utilized to identify distinct cell clusters. Inter-cluster similarity was validated via Pearson correlation. Bimod test (Wilcoxon rank-sum) was recruited to identify cluster-specific markers (FDR $<$ 0.05, $|\log FC| > 0.5$). Wilcoxon test with default parameters was performed to compare different sample groups (Submerged (S) vs Terrestrial (T), FDR $<$ 0.05, $|\log FC| > 0.5$) and identify differentially expressed genes (DEGs). Cluster visualization was generated by UMAP projections using the R package and uwot (v0.1.10). Cell type annotation was then performed to categorize cell populations. ClusterProfiler (v3.18.1) annotated differential genes with KEGG pathways (FDR $<$ 0.05). For statistical rigor, all tests used two-sided thresholds, and multiple testing was corrected via Benjamini-Hochberg.

Data Records

All the scRNA-seq raw data of *H. difformis* samples have been deposited to the NCBI Sequence Read Archive (SRA), with the number PRJNA1296768. The sample descriptions and raw sequencing data for scRNA-seq are available under the BioSample accession numbers (SAMN50199375 and SAMN50199374, respectively), and the Sequence Read Archive (SRA) accession number SRP603557 (SRR34716364 and SRR34716365, respectively)[32]. The processed transcriptomics data generated in this study have been deposited in the Gene Expression Omnibus (GEO) database under the accession number GSE309844[33].

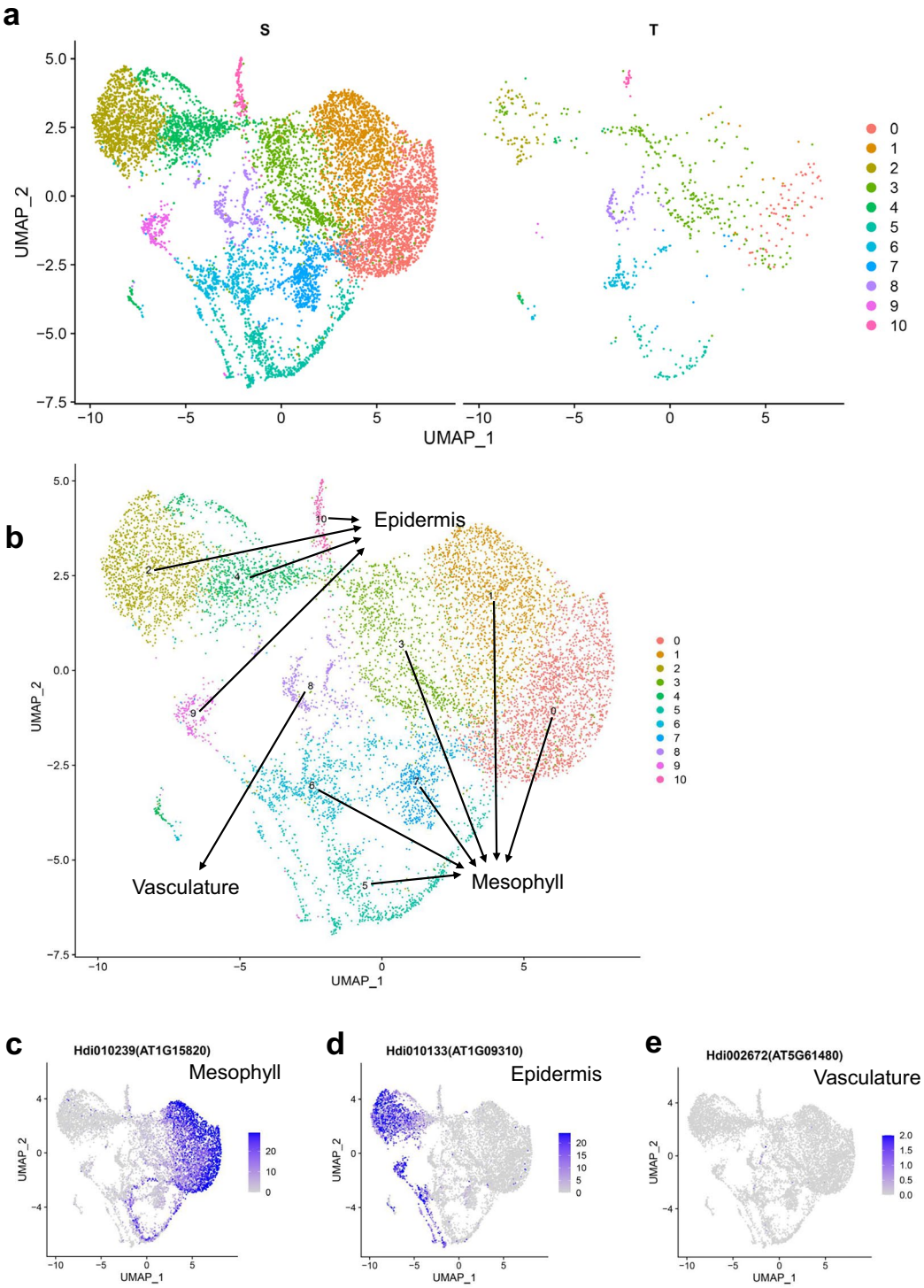

Fig. 3 Gene expression patterns of combined shoot cell clusters using scRNA-seq. (**a**) Uniform manifold approximation and projection (UMAP) plot showing the main cell types of eleven clusters in submerged (S) and terrestrial (T) samples. (**b**) UMAP plot showing the main cell types of emerged eleven clusters. (**c**) UMAP visualization of the typical mesophyll cell type-specific gene activity scores. (**d**) UMAP visualization of the typical epidermis cell type-specific gene activity scores. (**e**) UMAP visualization of the typical vasculature cell type-specific gene activity scores.

Data Overview

Using the 10 X Genomics scRNA-seq platform, we generated scRNA-seq datasets for terrestrial and submerged shoots, respectively. A total of 376,036,047 read pairs were obtained from the submerged shoots' scRNA-seq data, of which 91.10% had a quality score of 30. In contrast, a total of 368,518,926 read pairs were obtained

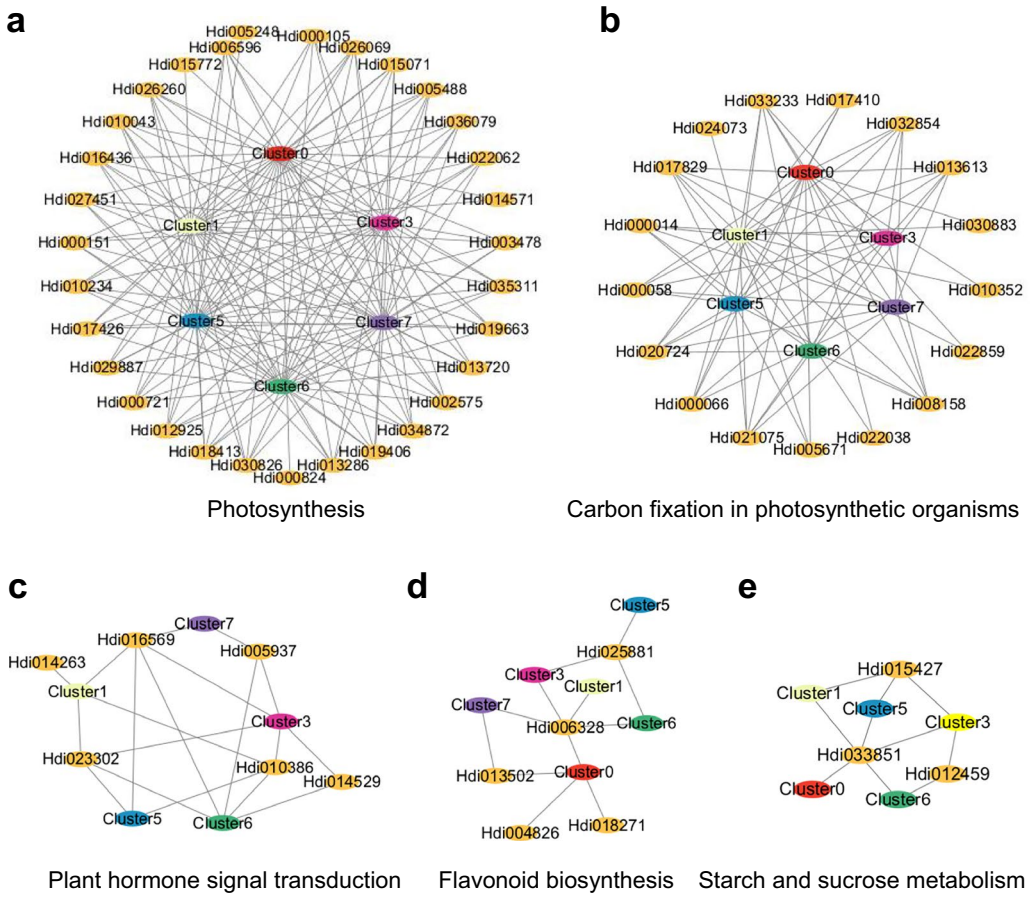

Fig. 4 KEGG annotated DEGs related to plant environmental adaptations in mesophyll cell clusters. (**a**) DEGs related to photosynthesis. (**b**) DEGs related to carbon fixation in photosynthetic organisms. (**c**) DEGs related to plant hormone signal transduction. (**d**) DEGs related to flavonoid biosynthesis. (**e**) DEGs related to starch and sucrose metabolism.

from the terrestrial shoots' scRNA-seq data, with 85.8% achieving a quality score of 30. The submerged shoots' sequencing read pairs were mapped to the *H. difformis* genome with a 91.5% mapping ratio, while terrestrial shoots' sequencing read pairs were mapped to the *H. difformis* genome with a 72.6% mapping ratio. At the single-cell resolution, a total of 9,361 submerged cells were obtained with a median number of 1,123 genes and an average of 40,170 reads per cell. In contrast, a total of 815 terrestrial cells were obtained with a median number of 424 genes and an average of 452,170 reads per cell. Overall, these results indicate that high-quality sequencing data were generated in the scRNA-seq datasets.

Technical Validation

Quality assessment of sequencing data. For quality control, we visualized the mitochondrial and chloroplast gene expression proportions, and the complexity defined by the ratio between number of genes detected (nGene) and number of unique molecular identifiers (nUMI) per cell in the scRNA-seq data. We filtered out low-quality cells, genes, and potential doublets, which left 9,653 cells (8,929 submerged sample cells and 724 terrestrial sample cells) and confirmed the expression patterns of the nUMI and nGene in the filtered scRNA-seq data (Fig. 2). The median values of submerged samples for nGene and nUMI were 1,114 and 3,472 respectively. The median values of terrestrial samples for nGene and nUMI were 452 and 682, respectively. Despite having fewer data points than the submerged sample, the results indicate that high-quality scRNA-seq data were successfully obtained to identify the cell-type specificity of shoots from *H. difformis*.

Gene expression analysis and cluster identification. After filtering all cells, we combined terrestrial and submerged cells and assigned them into eleven clusters using homology-based markers from *A. thaliana*. We constructed a total of eleven cell clusters[31]. Out of the eleven clusters analyzed, six clusters (clusters 0, 1, 3, 5, 6, and 7) were identified as mesophyll cells. Four clusters (clusters 2, 4, 9, and 10) were categorized as epidermal cells. The remaining cluster, cluster 8, was classified as vascular cells (Fig. 3a,b). We then identified homologs of marker genes which are specifically expressed in mesophyll cells, epidermal cells and vascular cells (Fig. 3c–e). These findings suggest that our scRNA-seq data effectively identified clusters by cell type.

KEGG annotation of differentially expressed genes. Furthermore, we identified differentially expressed genes (DEGs) between terrestrial and submerged samples in each cluster and annotated them by the Kyoto Encyclopedia of Genes and Genomes (KEGG) database[31]. In clusters representing mesophyll cells (clusters 0, 1, 3, 5, 6, 7), DEGs involved in plant adaptation to various environmental changes were enriched in the submerged sample (Fig. 4). These results indicate that our scRNA-seq data are effective for revealing the molecular mechanisms underlying environmental adaptation in aquatic plants.

Data availability

The raw data is available at the NCBI Sequence Read Archive (SRA) database (https://identifiers.org/ncbi/insdc.sra:SRP603557). The processed transcriptomics data generated in this study is available at the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE309844).

Code availability

All software used in this paper has been described in the Methods section with the respective version number. For this study, no custom code was generated.

Received: 11 August 2025; Accepted: 18 November 2025;
Published online: 29 November 2025

References

1. Li, G. *et al*. Water wisteria genome reveals environmental adaptation and heterophylly regulation in amphibious plants. *Plant Cell Environ* **47**, 4720–4740, https://doi.org/10.1111/pce.15050 (2024).
2. Xu, H., Li, W., Fu, W. L., Xie, Z. M. & Huang, W. M. Aquatic plant *Myriophyllum spicatum* displays contrasting morphological, photosynthetic, and transcriptomic responses between its aquatic and terrestrial morphotypes. *Photosynth Res* **163**, 15, https://doi.org/10.1007/s11120-025-01138-5 (2025).
3. Koga, H., Ikematsu, S. & Kimura, S. Diving into the water: Amphibious plants as a model for investigating plant adaptations to aquatic environments. *Annu Rev Plant Biol* **75**, 1–26, https://doi.org/10.1146/annurev-arplant-062923-024919 (2024).
4. van Veen, H. & Sasidharan, R. Shape shifting by amphibious plants in dynamic hydrological niches. *New Phytol* **229**, 79–84, https://doi.org/10.1111/nph.16347 (2021).
5. Wanke, D. The ABA-mediated switch between submersed and emersed life-styles in aquatic macrophytes. *J Plant Res* **124**, 467–475, https://doi.org/10.1007/s10265-011-0434-x (2011).
6. Lin, B. L. & Yang, W. J. Blue light and abscisic acid independently induce heterophyllous switch in *Marsilea quadrifolia*. *Plant Physiol* **119**, 429–434, https://doi.org/10.1104/pp.119.2.429 (1999).
7. Goliber, T. E. Endogenous abscisic acid content correlates with photon fluence rate and induced leaf morphology in *Hippuris vulgaris*. *Plant Physiol* **89**, 732–734, https://doi.org/10.1104/pp.89.3.732 (1989).
8. Nakayama, H. *et al*. Regulation of the KNOX-GA gene module induces heterophyllic alteration in North American lake cress. *Plant Cell* **26**, 4733–4748, https://doi.org/10.1105/tpc.114.130229 (2014).
9. Sato, M. *et al*. Temperature-dependent changes of cell shape during heterophyllous leaf formation in *Ludwigia arcuata* (Onagraceae). *Planta* **228**, 27–36, https://doi.org/10.1007/s00425-008-0715-3 (2008).
10. Kuwabara, A., Ikegami, K., Koshiba, T. & Nagata, T. Effects of ethylene and abscisic acid upon heterophylly in *Ludwigia arcuata* (Onagraceae). *Planta* **217**, 880–887, https://doi.org/10.1007/s00425-003-1062-z (2003).
11. Li, G. *et al*. Establishment of an Agrobacterium mediated transformation protocol for the detection of cytokinin in the heterophyllous plant *Hygrophila difformis* (Acanthaceae). *Plant Cell Rep* **39**, 737–750, https://doi.org/10.1007/s00299-020-02527-x (2020).
12. Nakayama, H., Sinha, N. R. & Kimura, S. How do plants and phytohormones accomplish heterophylly, leaf phenotypic plasticity, in response to environmental cues. *Front Plant Sci* **8**, 1717, https://doi.org/10.3389/fpls.2017.01717 (2017).
13. Zhao, W., Xiao, J., Lin, G., Peng, Q. & Chu, S. Morphological and physiological response of *amphibious Rotala* rotundifolia from emergent to submerged form. *J Plant Res* **137**, 279–291, https://doi.org/10.1007/s10265-024-01521-8 (2024).
14. Iida, S. *et al*. Loss of heterophylly in aquatic plants: not ABA-mediated stress but exogenous ABA treatment induces stomatal leaves in *Potamogeton perfoliatus*. *J Plant Res* **129**, 853–862, https://doi.org/10.1007/s10265-016-0844-x (2016).
15. Kim, J. *et al*. A molecular basis behind heterophylly in an amphibious plant, *Ranunculus trichophyllus*. *Plos Genet* **14**, e1007208, https://doi.org/10.1371/journal.pgen.1007208 (2018).
16. Koga, H., Kojima, M., Takebayashi, Y., Sakakibara, H. & Tsukaya, H. Identification of the unique molecular framework of heterophylly in the amphibious plant *Callitriche palustris* L. *Plant Cell* **10**, 3272–3292, https://doi.org/10.1093/plcell/koab192 (2021).
17. Han, S., Xing, Z., Jiang, H., Li, W. & Huang, W. Biological adaptive mechanisms displayed by a freshwater plant to live in aquatic and terrestrial environments. *Environ Exp Bot* **191**, 104623, https://doi.org/10.1016/j.envexpbot.2021.104623 (2021).
18. Ikematsu, S. *et al*. Rewiring of hormones and light response pathways underlies the inhibition of stomatal development in an amphibious plant *Rorippa aquatica* underwater. *Curr Biol* **33**, 543–556, https://doi.org/10.1016/j.cub.2022.12.064 (2023).
19. He, D. *et al*. Investigating the molecular basis for heterophylly in the aquatic plant *Potamogeton octandrus* (Potamogetonaceae) with comparative transcriptomics. *PeerJ* **6**, e4448, https://doi.org/10.7717/peerj.4448 (2018).
20. Li, G. *et al*. Water-Wisteria as an ideal plant to study heterophylly in higher aquatic plants. *Plant Cell Rep* **36**, 1225–1236, https://doi.org/10.1007/s00299-017-2148-6 (2017).
21. Kazmi, A. *et al*. Cadmium stress alters heterophylly, anatomy, and photosynthetic activity of *Hygrophila difformis*. *Ecotoxicol Environ Saf* **302**, 118685, https://doi.org/10.1016/j.ecoenv.2025.118685 (2025).
22. Horiguchi, G., Mizokami, Y., Hirotsu, N. & Noguchi, K. Characterization of biochemistry and anatomy for underwater carbon uptake in leaves of the amphibious plant *Hygrophila difformis*. *J Exp Bot* **76**, 4581–4597, https://doi.org/10.1093/jxb/eraf250 (2025).
23. Pal, D. & Samanta, K. CNS activities of ethanol extract of aerial parts of *Hygrophila difformis* in mice. *Acta Pol Pharm* **68**, 75–81 (2011).
24. Yi, R. *et al*. A single-cell transcriptomic dataset of pluripotent stem cell-derived astrocytes via *NFIB/SOX9* overexpression. *Sci Data* **11**, 987, https://doi.org/10.1038/s41597-024-03823-x (2024).
25. Kim, M. S. *et al*. Single-cell transcriptome sequencing of plant leaf expressing anti-HER2 VHH-FcK cancer therapeutic protein. *Sci Data* **10**, 911, https://doi.org/10.1038/s41597-023-02833-5 (2023).
26. Liang, X. *et al*. Single-cell transcriptomic analyses reveal cellular and molecular patterns of rubber tree response to early powdery mildew infection. *Plant Cell Environ* **46**, 2222–2237, https://doi.org/10.1111/pce.14585 (2023).
27. Zhu, M. Y. *et al*. Single-cell transcriptomics reveal how root tissues adapt to soil stress. *Nature* **642**, 721–729, https://doi.org/10.1038/s41586-025-08941-z (2025).

28. Liu, Z. *et al*. Integrated single-nucleus and spatial transcriptomics captures transitional states in soybean nodule maturation. *Nat Plants* **9**, 515–524, https://doi.org/10.1038/s41477-023-01387-z (2023).
29. Yue, H. *et al*. Single-cell transcriptome landscape elucidates the cellular and developmental responses to tomato chlorosis virus infection in tomato leaf. *Plant Cell Environ* **47**, 2660–2674, https://doi.org/10.1111/pce.14906 (2024).
30. Li, G. Genome annotation datasets of *Hygrophila difformis.V1*. *figshare* https://doi.org/10.6084/m9.figshare.25441387.v2 (2024).
31. Li, G. ScRNA-seq data of *Hygrophila difformis*. *figshare* https://doi.org/10.6084/m9.figshare.29632025.v2 (2025).
32. *NCBI Sequence Read Archive* https://identifiers.org/ncbi/insdc.sra:SRP603557 (2025).
33. *NCBI GEO* https://identifiers.org/geo/GSE309844 (2025).

Acknowledgements

This study was supported by the International Partnership Program of the Chinese Academy of Sciences (075GJHZ2024051FN; 075GJHZ2024072MI).

Author contributions

G.L. and H.H. contributed to research design, G.L. carried out laboratory analyses, bioinformatic analysis, visualized data obtained funding and wrote the original draft, A.K. and H.H. revised the manuscript. All authors approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.