The chromosome-level genome of Water-Wisteria establishes an ecological plant model and provides insight into the molecular mechanisms underlying heterophylly in amphibious plants

Gaojie Li[1], Xuyao Zhao[1], Jingjing Yang[1], Yimeng Chen[1,2], Yan Chen[1,2], Yixian Li[1,2], Zuoliang Sun[1,2], Manli Xia[1,2], Xiaoyu Wang[1,2], Xiaozhe Li[1], Kangsheng Luo[1], Shiqi Hu[3], Jathish Ponnu[4], Seisuke Kimura[5,6], Inhwan Hwang[7], Keiko U Torii[8,9,10], Hongwei Hou[1,2,]*

[1] The State Key Laboratory of Freshwater Ecology and Biotechnology, The Key Laboratory of Aquatic Biodiversity and Conservation of Chinese Academy of Sciences, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan, Hubei, 430072, China

[2] University of Chinese Academy of Sciences, Beijing 100049, China

[3] Zhejiang Marine Development Research Institute, Zhoushan, 316021, China

[4] Institute for Plant Sciences, University of Cologne, Zülpicherstraße 47b, Cologne D-50674, Germany

[5] Faculty of Life Sciences, Kyoto Sangyo University, Kamigamo-motoyama, Kita-ku, Kyoto-shi, Kyoto 603-8555, Japan

[6] Center for Plant Sciences, Kyoto Sangyo University, Kamigamo-motoyama, Kita-ku, Kyoto-shi, Kyoto 603-8555, Japan

[7] Department of Life Science, Pohang University of Science and Technology, Pohang 37673, Republic of Korea

[8] Institute of Transformative Biomolecules (WPI-ITbM), Nagoya University, Chikusa, Nagoya, Aichi, 464-8601, Japan

[9] Howard Hughes Medical Institute, The University of Texas at Austin, Austin, TX 78712, USA

[10] Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX 78712, USA

*Corresponding Author: Email: houhw@ihb.ac.cn (H. H.)

Summary

The phenomenon by which a single plant dramatically changes its leaf shape in response to the surrounding environment is called heterophylly, an ideal model for the study of plant phenotypic adaptation to environment. *Hygrophila difformis* (Acanthaceae), also known as Water-Wisteria, has recently merged as a model to study heterophylly due to its typical phenotypic plasticity to various ecological factors. Here, we generated a chromosome-level genome of *H. difformis* (scaffold N50: 57.63Mb, and genome size: 831.53 Mb) and predicted 36,099 protein-coding genes mainly on 15 pseudochromosomes. *H. difformis* evolved a gene expansion for its acclimation to the amphibious environment. RNA-seq analysis revealed that genes relative to environmental stimuli, leaf development and other processes are elaborated, regulating its morphological and physiological acclimations to the changing environments. Finally, we identified candidate genes under different ecological factors, and illuminated the roles of *LATE MERISTEM IDENTITY 1* (*LMI1*) on the regulation of heterophylly in *H. difformis*. Our study contributes the first genome and transcriptome for *Hygrophila*, establishing *H. difformis* as a model for the relationships of ecological factors and plant morphological features. We provide new insights into the adaptability of plants in evolution, along with potential resources for trait improvement in crops.

Keywords: Heterophylly, Environmental adaptation, Genome, *Hygrophila difformis*, Molecular mechanisms

Introduction

Ancestral angiosperms returned to the water millions of years ago and evolved into aquatic plants (Wissler et al., 2011). Amphibious plants are those that can grow in both terrestrial and submerged conditions and evolved unique advantages between different environments (Agarie et al., 1997). Profound investigations have been performed on terrestrial plants, and a few genomic data of aquatic plants were also recently reported (Lu et al., 2021; Shi et al., 2022). However, genomic data of amphibious plants are still lacking. Heterophylly, the distinct morphological plasticity in leaf form in response to environmental changes within a single plant, was widely seen in amphibious plants (van Veen and Sasidharan, 2021). Heterophylly is manifested in response to various environmental factors, such as light density and quality (Momokawa et al., 2011; Nakayama et al., 2014), temperature (Nakayama and Kimura, 2015), and surrounding gases (Kuwabara et al., 2003). Notably, heterophyllous plants have evolved leaf variation for morphological acclimations across diverse taxa (Fig. S1), thereby providing ideal models of plant environmental adaptation and convergent evolution, as well as the corresponding relationships between plant morphological features and ecological factors. Recently, mechanisms regulating heterophylly have been studied in some species. For instance, ethylene and abscisic acid (ABA) regulate genes that are involved in the adaxial–abaxial polarity and induce heterophylly in *Ranunculus trichophyllus* (Ranunculaceae) (Kim et al., 2018). Genes related to cell elongation and division contribute to the leaf differentiation in *Callitriche palustris* (Plantaginaceae) (Koga et al., 2021). In *Rorippa aquatica* (Brassicaceae), the *KNOTTED1-LIKE homeobox* (*KNOX1*) genes regulate the accumulation of endogenous phytohormones gibberellin (GA) and cytokinin (CK), to regulate the temperature-dependent heterophylly (Nakayama et al., 2014). However, even though a few transcriptomic data of several heterophyllous plants have been reported (He et al., 2018; Koga et al., 2021; van Veen et al., 2013), their molecular mechanisms are still largely unknown.

Hygrophila difformis, also known as Water-Wisteria, is an amphibious plant which develops deeply lobed, complex leaf forms in submerged conditions but simple

leaves in terrestrial conditions (Fig. 1a). *H. difformis* is a highly promising model to investigate the mechanism of heterophylly as it recently gained the attention of researchers for its advantages such as easy gene transformation, fast propagation, and conspicuous heterophylly in response to environmental stimuli (Li et al., 2017; Li et al., 2020). Moreover, it provides a suitable ecological model for explicitly addressing the effects on environmental responses, morphological features, and gene expression variations in an ecological context (Li et al., 2021). *Hygrophila* contains almost 90 species, and has a large number of amphibious species which developed phenotypic plasticity (Fig. S2), among which *H. difformis* showed the most conspicuous heterophylly. Most species of *Hygrophila* are famous as ornamental plants (Karatas et al., 2013), useful as medicinal herbs (Ingale et al., 2013), and also serve as models for photosynthetic studies (Botts et al., 1990; Horiguchi et al., 2021). However, at present there is no available genomic resources in this genus.

Previously, many studies on model plants have contributed to understanding the role of gene function in plant phenotypes. It remains unclear how the same genotype regulates phenotypic plasticity in different environments. Here, we report the chromosome-scale genome of *H. difformis*, by combining PacBio sequencing, Illumina sequencing and High-throughput Chromosome Conformation Capture (Hi-C) sequencing technology. Genomic analyses reveal that gene expansion occurred during *H. difformis* speciation, and the expansion of transcription factors (TFs) may contribute to its amphibious habitat. Transcriptome analyses based on our genomic data reveal molecular mechanisms for its morphological and physiological acclimations. We also identified candidate genes *LATE MERISTEM IDENTITY 1* (*LMI1*) and verified their roles by genetic methods. Our study contributes the first genome and transcriptome for *Hygrophila*, establishing *H. difformis* as a model for the understanding of environmental functions on plant phenotypes.

Results and discussion

Genome sequencing, assembly, and annotation

The genome size of *H. difformis* was estimated to be 913.38 Mb by the genomic survey and *K*-mer analysis (Fig. S3; Table S1). The basic number of chromosomes in

the related *Hygrophila* species is 16 and diploid species show 2n = 2x = 32 (Sharma et al., 2013). Our cytogenetic studies showed 45 chromosomes, indicating that the sequenced *H. difformis* individual may be a triploid (2n= 3x= 45) (Fig. S4). The fact that the *H. difformis* chromosome number is not a multiple of 16 indicates that its genome may be a result of chromosome fusion need to be further investigated. By using PacBio Sequel platform (Zhou et al., 2021), the genome of *H. difformis* was sequenced, separated and assembled to the haploid genome. This initial genome assembly comprised 2,752 contigs covering 871.90 Mb, with a contig N50 of 0.68 Mb. Using the High-throughput Chromosome Conformation Capture sequencing (Hi-C seq) (Burton et al., 2013; Durand et al., 2016) platform, 95.37% (831.53 Mb) contig sequences have been anchored to 15 pseudo-chromosomes, with a contig N50 of 0.67 Mb (Table 1; Table S2-4). Based on the heatmap of the contig contact matrix with Hi-C data (Fig. 1b), we concluded that the clustering, ordering and orientation of the contigs were good. According to homology-based and de novo predictions, 36,099 protein-coding gene models were identified (Table S5), and 90.62% (32,714) of the predicted protein-coding genes could be mapped with functional annotations using public databases (Table S6). Through a combination of homology-based searches and de novo prediction, we identified repetitive sequences accounting for 74.92% of its genome (Fig. 1c; Table S7), more than its terrestrial related species *Andrographis paniculata* (53.3%) (Sun et al., 2019) and its aquatic related species *Utricularia gibba* (42.18%) (Lan et al., 2017). Among them, long terminal repeat (LTR) were the dominant repeat sequences (62.39%) of the assembled genome (Table S7). BUSCO (Simao et al., 2015) (Benchmarking Universal Single-Copy Orthologs) analysis suggests that 95.50% of genes were recovered (Table S8). These results indicated successful assembly of the major genic regions of the *H. difformis* genome.

Comparative genomics and genome evolution

To investigate the evolution of *H. difformis*, we compared its genome to fourteen representative species of diverse evolutionary statuses, including one monocot - *Oryza sativa* (Poaceae) and thirteen eudicots. A total of 14,656 gene families were identified, of which 1,362 were species-specific in *H. difformis*. *H. difformis* shared more

common gene families with *O. europaea* (11,678), *B. hygrometrica* (10,099), and *E. guttata* (10,983) (Figs 2c, 2d; Table S9). 210 single-copy orthologous gene sets were identified among these 15 genomes, assuming the molecular clock (Kumar et al., 2017) to construct a species phylogenetic tree. Among Lamiales, we found the estimated divergence times were around 69.8 million years ago (Mya) between *O. europaea* and the other six species, 57.8 Mya between *B. hygrometrica* and the remaining five species, and 30.6 Mya between *A. paniculata* and *H. difformis* (Fig. 2a).The phylogenetic placement of these species was consistent with the previous studies (Kordyum and Mosyakin, 2020).

We then investigated the gene family expansion and contraction in the phylogenetic analysis. We found genes in *Glycine max*, *H. difformis* and *O. europaea* were highly expanded, and *H. difformis* showed a significant gene expansion than its closely related species (Fig. 2b). A total of 5,089 gene families were expanded, while 1,390 gene families were contracted in the *H. difformis* lineage. We then tested for enrichment of gene ontology (GO) terms by comparing the annotation content of expanded or contracted genes in *H. difformis*. Notably, the top enriched GO pathways among expanded gene families was response to wounding (GO:0009611). As vegetative propagation is known to be predominant in aquatic plant (Amano et al., 2020), and *H. difformis* has easy regeneration through the wound induced vegetative propagation (Li et al., 2017), the expansion of these gene families may be involved in its vegetative propagation under aquatic conditions. Genes related to environmental stimuli were also identified in the expanded families, including response to temperature (GO:0009266), osmotic stress (GO:0006970), blue light (GO:0009637) and phototropism (GO:0009638), which may involve in its responses to diverse environments.

Whole genome duplication (WGD) events are commonly regarded as a major evolutionary force, which happened frequently in plants (Ren et al., 2022). The synonymous substitution divergence (*Ks*) of each collinear gene pair within a genome or between genomes was then characterized for WGD analysis between *H. difformis* and other related species (Fig. 2e). The *Ks* distribution of paralogs within *H. difformis*

and *A. paniculata* suggested only one peak at *Ks* ~1.19 and ~1.41, while *S. indicum* has two peaks at *Ks* ~0.62 and ~1.69. The orthologous peaks for *H. difformis*– *S. indicum* and – *A. paniculata*, correspond to *Ks* values of ~0.76 and ~0.73, respectively. These results imply that the *H. difformis* genome may have experienced only one ancient whole-genome triplication (γ), and no specific WGD event occurred after the divergence from *A. paniculata*. To infer intra-genomic gene collinearity, we identified a total of 560 homologous blocks involving 11,281 collinear gene pairs in *H. difformis*. The largest homologous block with 108 gene pairs, was located in chromosomes 7 and 9. To infer inter-genomic gene collinearity, we identified a total of 1,081 homologous blocks involving 30,551 collinear gene pairs between the *H. difformis* and *A. paniculata* genome. The largest homologous block in chromosome 8 of *H. difformis* and chromosome 18 of *A. paniculata*, with 361 gene pairs (Fig. S5).

Expansion of transcription factors (TFs) contributed to environmental adaptation in *H. difformis*

Transcription factors (TFs) are those that interact with specific sequences via DNA-binding domains, crucial for transcriptional regulation (Lehti-Shiu et al., 2017). To identify the expansion in TFs, we initially conducted a comprehensive characterization of *H. difformis*, compared to other eight aquatic plants (*Azolla filiculoides*, *Nymphaea colorata*, *Spirodela polyrhiza*, *Zostera marina*, *Ceratophyllum demersum*, *Nelumbo nucifera*, *Aldrovanda vesiculosa* and *U. gibba*) by PlantTFDB (Jin et al., 2017). Among the reported TF families, seventeen were identified to be highly expanded in *H. difformis*, including ARR-B, B3, bHLH, bZIP, C2H2, CAMTA, ERF, GRAS, HSF, LBD, M-type_MADS, MYB_related, NAC, Nin-like, WOX, WRKY and ZF-HD (Fig. 2f). Among those, NAC and WRKY families play important roles in plant responses to the environmental stress, such as drought (Mao et al., 2015), heat (Cheng et al., 2021b), and oxidative stress (Yuan et al., 2019), which may devote to the acclimation of *H. difformis* in the amphibious surroundings. B-type RESPONSE REGULATOR (ARR-B) TFs were reported to be related to the response to CK (Sakai et al., 2001), while the B3 family and GRAS family are known in the response to the auxin (Swaminathan et al., 2008) and GA (Cenci and Rouard, 2017).

ETHYLENE RESPONSE FACTORs (ERFs) are involved in ethylene signaling (Debbarma et al., 2019). For the reason that phytohormones are essential to environmental acclimation and plant development (Li et al., 2021), these expanded TFs may contribute to the superior adaptation of *H. difformis* to survive in dehydrated environments compared to the other aquatic plants.

We then compared TFs of *H. difformis* with its closely related terrestrial species (*A. paniculata*, *E. guttata*, *B. hygrometrica*, *G. aurea* and *S. indicum*) and found eleven expanded TFs families in *H. difformis*, including BES1, ERF, HB-other, HSF, MYB_related, NF-X1, SRS, TCP, WOX, YABBY and ZF-HD (Fig. 2g). Compared to the closely related species, five shared expanded TF families (ERF, HSF, MYB_related, WOX, and ZF-HD) were found in *H. difformis*. Heat shock factors (HSFs) play important roles during plant growth in response to heat and many other environmental stresses (Yu et al., 2022), and are responsible for the acclimation to the rapidly changing environments (Huang et al., 2018). Zinc finger-homeodomain (ZF-HD) proteins also function in plant growth, and phytohormone biosynthesis, and are involved in the responses to abiotic and biotic stresses (Lai et al., 2021). It is well-known that TCPs is essential in the leaf shape determination (Koyama et al., 2010; Yu et al., 2021). To explore the classification of TCPs in *H. difformis*, we constructed a phylogenetic tree using 87 TCP proteins of *H. difformis*, *A. paniculata* and *A. thaliana*. After further identification, we found that all TCPs can be clarified into three classical groups: the CIN family, CYC/TB1 family, and the PCF family. The copy numbers of TCPs from *H. difformis* in each group are highly increased than in *A. paniculata*, as the CIN family (18:10), CYC/TB1 family (7:3), and the PCF family (16:9) (Fig. 2h). These redundant genetic resources may provide the genetic basis for the diversification of leaf plasticity in *H. difformis*. In addition, WUSCHEL-related homeobox (WOX) genes are also involved in leaf development and they regulates the expression of CIN class TCPs (Jha et al., 2020; Wang et al., 2020). Many aquatic species have evolved to be polyploid during species differentiation (Albert and Renner, 2020; Lu et al., 2021; Prancl et al., 2014), and polyploid was verified to be one of the strategies for plants to expand their ranges and invade new environments by endowing

polyploid progenies with novel traits (Cheng et al., 2021a). In addition, gene expansion (especially TFs) is important for plant evolution (Feller et al., 2011; Mathew and Agarwal, 2018). Therefore, these expanded TFs may contribute to the specific phenotypic innovation and acclimation in *H. difformis*.

Morphological, physiological and molecular adaptation of *H. difformis* to the terrestrial/submerged environment

The heterophyllous plants grown in submerged conditions always developed filiform, dissected or linear leaves, lacking cuticles and stomata; while terrestrial leaves are thicker, and have differentiated mesophylls and cutinized with stomata (Li et al., 2021). Here we found that both leaf area, dissection index (DI) and epidermal cell complexity of *H. difformis* were significantly upregulated in submerged conditions, while stomatal and vein density, and chlorophyll content were significantly upregulated in terrestrial conditions (Fig. 3a, S6), which is similar to the characteristics in other heterophyllous plants. To study the underlying molecular mechanisms, we performed RNA-seq analysis between submerged and terrestrial shoots and noted them based on our genomic data. A total of 6,226 differentially expressed genes (DEGs) (false discovery rate [FDR] < 0.05 and expression-level $\log_2$ [fold-change] > 1) were identified, with 4,100 upregulated genes and 2,126 downregulated genes in terrestrial shoots. Using GO enrichment, we identified representative upregulated genes in terrestrial conditions (Fig. 3b). We found that genes “response to oxidative stress” have higher count numbers among the response, which is consistent with our discovery that reactive oxygen species (ROS) and their related genes are accumulated in terrestrial leaves of *H. difformis* (Figs 3d, S7). We also observed upregulation of genes belonging to the category “Xylan”, suggesting their important roles in the terrestrial adaptation of *H. difformis*. We found upregulated genes in the submerged shoots involved in “Response to osmotic stress” and “Cell wall organization” (Fig. 3c), which may contribute to the submerged acclimation of *H. difformis*.

We then analyzed the DEGs that are involved in environmental sensing or responses (Fig. 3d). The thermosensor gene *EARLY FLOWERING 3* (*ELF3*) (Jung et

al., 2020), which is a negative regulator of thermomorphogenesis, is significantly downregulated in terrestrial conditions (Fig. 3d). For the reason that *ELF3* functions antagonistically with *PHYTOCHROME-INTERACTING FACTOR4* (*PIF4*) in thermomorphogenesis (Raschke et al., 2015), we also found their conversed expression pattern in *H. difformis*. The proton sensor *SLAC1 HOMOLOGUE 3* (*SLAH3*) (Lehmann et al., 2021), the mechanical stimuli sensor *MSCS-LIKE 10* (*MSL10*) (Tran et al., 2021), and aquaporins *Plasma membrane Intrinsic Proteins* (*PIPs*) (Yaneff et al., 2015) are also significantly upregulated in terrestrial conditions.

Morphological changes in leaf shape are the most obvious characteristics in heterophyllous plants. It is known that *KNOX1* and *BEL1-like homeobox* (*BELL* or *BLH*) functions in an antagonistic manner to regulate leaf development, as the overexpression of *KNOX1* increases leaf complexity (Hareven et al., 1996) whereas increasing of the *BLH* genes causes leaf simplification (Kimura et al., 2008; Kumar et al., 2007). We found that *BELL* and *BLH* are significantly upregulated in terrestrial shoots, which may lead to leaf simplification in terrestrial conditions. Overexpression of *TfTCP13* (from *Torenia fournieri*) in *A. thaliana* displayed narrower leaves (Zhang et al., 2021), and we found *TCP13* is upregulated in submerged shoots of *H. difformis*. *TCP14* is downregulated in terrestrial leaves of *H. difformis*, similar to the broad, short leaf form caused by *TCP14::SRDX* in *A. thaliana* (Kieffer et al., 2011). *GROWTH-REGULATING FACTORs* (*GRFs*) regulate leaf size in *A. thaliana* (Debernardi et al., 2014). *GRF5* is upregulated in the submerged shoots of *H. difformis*, which may contribute to its increasing leaf size in submerged conditions (Fig. 3a). In addition, genes involved in chloroplast development and chlorophyll biosynthesis are upregulated in terrestrial shoots (Figs 3a, S6), suggesting their functions in the photosynthetic plasticity under terrestrial or submerged conditions.

MUTE, *FAMA*, and *SLOW ANION CHANNEL-ASSOCIATED 1* (*SLAC1*), which play roles in stomata development and movement are significantly upregulated in terrestrial shoots, which may contribute to the stomatal formation in terrestrial leaves. It is worth mentioning that *EPIDERMAL PATTERNING FACTOR-LIKE9* (*EPFL9*/*STOMAGEN*) is a positive regulator of stomatal development (Hunt et al.,

2010), which is commonly absent in aquatic plants, including *Z. marina*, *S. polyrhiza*, *N. colorata*, and *W. Australiana* (Park et al., 2021). Here we identified its putative homologous gene (Hdi026079) in *H. difformis*. However, its expression is not detected in the terrestrial or submerged shoot, indicating that this gene may be a pseudogene or it evolved a functional differentiation in *H. difformis*. Genes involved in vascular development, lignin biosynthesis and wax biosynthesis are significantly upregulated in terrestrial shoots (Fig. 3b), which may lead to the terrestrial acclimation. A recent study in *Welwitschia* has shown that the DNA methylome maintains the link between gene expression, leaf development and ecological adaptation (Wan et al., 2021), and we found genes related to methylation are also changed in different conditions (Fig. 3b).

Roles of auxin and ABA in terrestrial or submerged acclimation

The theory that the morphogenesis induced by the environment is achieved by phytohormones is summarized by recent reviewers (Li et al., 2019; Nakayama et al., 2017). After counting DEGs involved in “Response to hormone”, we identified a large number of genes related to “Response to auxin” and “Response to abscisic acid” (Fig. 3g). Our previous study has verified that exogenous ABA can simplify the leaf shape of *H. difformis* (Li et al., 2017). Here, we found that genes related to ABA biosynthesis (e.g. *NINE-CIS-EPOXYCAROTENOID DIOXYGENASE* (*NCED*), *ABA DEFICIENT1* (*ABA1*), *ALDEHYDE OXIDASE* (*AAO*)) and ABA signaling-related genes are highly upregulated in terrestrial shoots, consistent with the increased ABA content in the terrestrial shoots (Fig. 3e, h). *CYP707A*, which encode ABA-degradation enzyme is downregulated in terrestrial conditions, similar to the downregulation of *PYRABACTIN RESISTANCE 1-LIKE* (*PYL*) genes. The expression pattern of these genes is consistent with their orthologs of *C. palustris* (Koga et al., 2021), implying their similar ABA pathways in the regulation of heterophylly.

Subsequently, we found the auxin biosynthesis gene, *YUCCA* (*YUC*) and the *YUC*-related *NAKED PINS IN YUC MUTANTS 1* (*NPY1*) are upregulated in submerged shoots, consistent with the increased IAA content in the submerged shoots (Fig. 3f, h). *INDOLE-3-ACETIC ACID INDUCIBLE* (*IAA*) and *AUXIN RESPONSE*

FACTOR 10 (*ARF10*) are also upregulated in submerged conditions. Interestingly, the auxin influx transporter *AUXIN RESISTANT 1* (*AUX1*) is upregulated in submerged conditions, while the auxin efflux carrier *PIN-FORMED* (*PIN*) and *PIN-LIKES* (*PILS*) are downregulated in submerged conditions. To further investigate the effects of auxin transport in *H. difformis*, we grew submerged plants in solutions containing the polar auxin transport inhibitor naphthylphthalamic acid (NPA), and NPA induced a gradual reduction of leaf dissection and ultimate formation of an oval leaf form with a smooth leaf margin (Fig. 3i). Previous studies have reported the functions of ABA and ethylene (Kuwabara et al., 2003), GA and CK (Nakayama et al., 2014) in heterophylly. Here, we report for the first time that auxin metabolism and auxin transport are also involved in the regulation of heterophylly in *H. difformis*.

Expression patterns of DEGs under diverse environmental conditions

A previous study has revealed expression patterns of candidate genes in the heterophyllous plant *C. palustris* under different phytohormone treatments (Koga et al., 2021). However, it is still unclear how heterophyllous plants respond to the natural, diverse environmental signals. Based on the typical phenotype of *H. difformis*, these plants were divided into four ecological groups (submerged/terrestrial, normal/low light density, normal/low temperature, and high/normal humidity) (Fig. 4a). After transcriptomic analysis, we identified a total of 6,226 DEGs in the group1 (Fig. 3); 9,213 DEGs in the group2; 2,851 DEGs in the group3; and 4,065 DEGs in the group4 (Figs 4b, 4c). A total of 180 genes were identified as common upregulated DEGs and 119 genes were identified as common downregulated DEGs among the four comparisons (Table S10-11). *NCED3*, *ETHYLENE-INSENSITIVE 3* (*EIN3*), *CYTOKININ OXIDASE 5* (*CKX5*), and *GA INSENSITIVE DWARF1A* (*GID1A*) genes, which were related to phytohormone pathways are significantly upregulated in all groups (Fig. 4d). Genes related to antioxidants are also significantly upregulated in all groups. *BETA CARBONIC ANHYDRASE 4* (*BCA4*) is upregulated, while *NPY1* is downregulated in all groups (Fig. 4e). A previous study has shown that *SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTERS* (*SWEET*) genes contributed to the specific development strategy of *Euryale ferox* leaves through expression and

function divergence (Wu et al., 2022). Here, we found that three orthologs of *SWEETs* are upregulated in all groups, implying their potential roles in heterophylly (Fig. 4d). We also identified nineteen upregulated TFs and eight downregulated TFs (Fig. 4f) from these common DEGs through PlantTFDB, which may regulate the acclimation of *H. difformis* under different ecological factors.

To further identify key genes involved in the heterophylly of these groups, co-expression networks were performed on a total of 15,295 DEGs (all DEGs that were significantly changed in at least one group). A total of 8,000 genes from 15 samples were ultimately clustered into ten modules using weighted gene co-expression network analysis (WGCNA) (Fig. S8). After analysis, we focused on the yellow module, which was significantly associated with gene expression pattern and leaf complexity (Fig. S9). We identified three leaf development-related genes: orthologs of *LATE MERISTEM IDENTITY1* (*LMI1A*, *LMI1B* and *LMI1C*) are positively co-expressed with three auxin-related genes (*NPY1*, *AUX1* and *ARF16*), two *GA-STIMULATED IN ARABIDOPSIS* (*GASA7* and *GASA14*), one ethylene synthesis gene *ACC SYNTHASE1* (*ACS1*), and a photosynthesis-related gene *LIGHT-HARVESTING COMPLEX B7* (*LHCB7*) (Fig. 4g). These candidate genes involved in heterophylly may regulate the complex leaf formation, together with phytohormone pathways and physiological adaptation to environmental changes.

HdLMI1 genes regulate the leaf shape of _H. difformis_ in submerged condition

Studies have been performed on *LMI1* and its orthologs in some terrestrial species, confirmed their conserved roles in the regulation of leaf shape diversity (Andres et al., 2017; Vlad et al., 2014). However, it is still unknown whether they play roles in aquatic plants. Firstly, we identified orthologs of *AtLMI1* in ten aquatic species, *H. difformis* and its terrestrial-related plant *A. paniculata*. Among those, we found that *LMI1* seems to be an innovation of eudicots, as they were not identified from the ferns (*A. filiculoides*), monocots (*S. polyrhiza* and *Z. marina*), basal angiosperms (*N. colorata* and *E. ferox*) and Ceratophyllales (*C. demersum*). In total, we obtained 2 orthologs in *N. nucifera*, *C. demersum*, *C. chinensis* and *A. paniculata*; 3 orthologs in *H. difformis*; 6 orthologs in *A. vesiculosa*; and 9 orthologs in *U. gibba*

(Fig. 5a; Table S7). Leaves vary from flat laminas to elaborate cup-shaped traps and trap formation can be explained by the gene expression pattern changes (Whitewoods et al., 2020). Since that *A. vesiculosa* and *U. gibba* are well-known carnivorous plants which evolved specialized, complex traps for trapping prey, the expansion of *LMI1s* may involve in their trap development and evolution.

To further investigate the functions of *HdLMI1s* in *H. difformis*, We knocked down these genes using the VIGS technique (Zhang et al., 2020) with Tobacco rattle virus (TRV) constructs, TRV2-HdLMI1A, TRV2-HdLMI1B and TRV2-HdLMI1C. Compared with the mock plants, the TRV2-HdLMI1A, TRV2-HdLMI1B and TRV2-HdLMI1C transgenic plants grown in submerged conditions with strong phenotypic changes were very similar to each other, which developed crinkled, complete or shallow serrated leaf margin (Fig. 5b). DI indicated that the mock plants have higher leaf complexity than TRV2-HdLMI1A, TRV2-HdLMI1B and TRV2-HdLMI1C transgenic plants (Fig. 5c). Following their morphological features, significantly decreased expression of all three *HdLMI1s* were detected in these transgenic plants (Fig. 5d), which is consistent with their co-expression patterns (Fig. 4g). A previous study has shown that orthologs of *LMI1* in *Medicago truncatula* involve in physical interactions and potentially form heterodimers, which cooperatively regulate leaf margin development (Wang et al., 2021). These results confirmed our prediction that *HdLMI1s* play key roles in the regulation of heterophylly of *H. difformis* in submerged conditions.

In a rapidly changing climate, flooding stress (submergence or waterlogging) is a major threat for many terrestrial plants, especially cash crops (Bailey-Serres et al., 2012). Plants have evolved two main strategies to overcome the lethal effects of submergence (Akman et al., 2012; Bailey-Serres and Voesenek, 2008; Muller et al., 2021). For the “Escape” strategy, plants elongate their shoot in an attempt to reach the surface and restore air contact. In contrast, the plants that use the “Wait” strategy reduce the levels of growth, and conserve energy reserves and delay the energy crisis, waiting for the flood to recede. Recent studies have mainly focused on the “Escape” strategy in crops, especially in deep-water rice (Kuroha et al., 2018; Nagai et al.,

2020). Interestingly, it was recently reported that the aquatic plant *N. nucifera* also utilized the "Escape" strategy for early submergence (Deng et al., 2022). Unlike the "Wait" or "Escape" strategy, heterophyllous plants seem to have evolved a third strategy - "Variation", to adapt to the constantly changing submergence or dehydration (Fig. 6). Our published genome of *H. difformis* provides resources and from which further studies will be possible to broaden our knowledge of the adaptations that have enabled the plant to survive in the fluctuant environment. Understanding the mechanisms of heterophylly will be a novel strategy for engineering resilient crops or agricultural development associated with climate change.

Material and Methods

Plant materials and phenotypic analysis

The sequenced individual of *H. difformis* was collected from South China Botanical Garden, Chinese Academy of Sciences, Guangdong Province, China (N113.373781º, E23.18794º). Using tissue culture, we raised plantlets from this individual to facilitate isolation of the high molecular weight genomic DNA. We also collected a large range of tissue samples (stem, shoot, root, flower and leaves) from the same plant for RNA sequencing (RNA-seq) for helping with assembly and assessment. Genomic DNA was isolated using the Qiagen DNeasy Plant Mini Kit (Qiagen, USA). For different environmental treatments, plant shoots were harvested from plants grown under different treatment (A: terrestrial plants grown at 23°C, 60% RH, 60 μmol m^{-2} s^{-1} illumination; B: terrestrial plants grown at 18°C, 60% RH, 60 μmol m^{-2} s^{-1} illumination; C: submerged plants grown at 23°C, 60 μmol m^{-2} s^{-1} illumination, D: submerged plants grown at 23°C, 10 μmol m^{-2} s^{-1} illumination, E: terrestrial plants grown at 23°C, 30% RH, 60 μmol m^{-2} s^{-1} illumination). Total RNA was isolated using TRIzol (Invitrogen, USA) for further studies. For morphological analysis, leaves were photographed with a Canon EOS80D camera, and all light microscopy observations were performed under a Sunny EX20 light microscope and photographed with a ToupCam TP605100A digital camera. Images were integrated using MvImage media software (ToupCam). Leaf complexity was estimated based on

the dissection index (DI), calculated as $\mathrm{DI} = \mathrm{leaf\ perimeter}/\sqrt{\mathrm{leaf\ area}}$, and epidermal cell complexity was also quantified by the same method as $\mathrm{Cell\ complexity} = \mathrm{cell\ perimeter}/\sqrt{\mathrm{cell\ area}}$. All calculations were performed using ImageJ 1.47v (http://rsb.info.nih.gov/ij/). Statistical differences were determined using Student's *t*-test, and data are results of twenty independent individuals.

Estimation of genome size and ploidy

The short reads from the MGIseq2000 platform were filtered and used for genome size estimation. The genome size was estimated by 17-*K*-mer analysis based on the *K*-mer method. The sequenced individual of *H. difformis* was determined with 45 chromosomes, using a cytological method described previously (Li et al., 2017).

De novo sequencing and genome assembly

Two long- insert continuous long reads libraries (40 kb) were prepared following PacBio Sequel (Pacific Biosciences of California) instructions. The SMRT (Single Molecule Real-Time) library was constructed using the SMRTbell Express Template Prep kit 2.0 (Pacific Biosciences). The library quantification and size were measured using Qubit 3.0 Fluorometer (Life Technologies) and Agilent 2100 system (Agilent Technologies). Subsequently, sequencing was performed on a PacBio Sequel II platform. With the PacBio Sequel platform, we filtered polymerase reads and generated subreads by removing adaptor sequences within sequences. The draft assembly of the genome was then assembled using MECAT2 (v20190226) (Xiao et al., 2017).

To anchor hybrid scaffolds onto chromosomes, genomic DNA was extracted from the leaves of one individual to construct a Hi-C library. We obtained sequencing data using an BGI MGISEQ-2000 platform with 150 PE mode. For anchored contigs, 381,400,686 clean read pairs were generated from the Hi-C library and mapped to the polished *H. difformis* genome using BWA (bwa-0.7.16) (Li and Durbin, 2010). Lachesis (Burton et al., 2013) was further applied to order and orient the clustered contigs. Juicer and Juicebox tool (v1.8.8) (Durand et al., 2016) were used to correct the assembly error in the Hi-C assembled genome.

Prediction and annotation of protein-coding genes

We predicted the protein-coding genes of *H. difformis* using three methods, including ab initio, homology-based and transcriptomic prediction. AUGUSTUS (v3.3.1) (Stanke et al., 2006) and GlimmerHMM (v3.0.4) (Majoros et al., 2004) were used to perform ab initio gene prediction. For the homology-based prediction, we used EXONERATE (v2.2.0) (Slater and Birney, 2005) to conduct gene prediction. For the transcriptomic prediction, we used RNA-seq to conduct gene prediction. Finally, MAKER (v3.00) (Cantarel et al., 2008) was used to integrate the prediction results. Gene functions were inferred according to the best match of the alignments to the nonredundant database (NR) of the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/), TrEMBL (Boeckmann et al., 2003), SWISS-PROT (Boeckmann et al., 2003), and InterPro (Mitchell et al., 2015). Motifs and protein domains were annotated using PfamScan (Mistry et al., 2007) and InterProScan (http://www.ebi.ac.uk/InterProScan). We also mapped the predicted protein-coding genes to their corresponding KEGG pathways, with the best hits for each gene used to establish GO and KEGG pathways.

Gene family identification and phylogenetic analysis

We selected 14 species to construct putative gene families, including *Oryza sativa* (Poaceae), *Boea hygrometrica* (Gesneriaceae), *Erythranthe guttata* (Phrymaceae), *Olea europaea* (Oleaceae), *Sesamum indicum* (Pedaliaceae), *Andrographis paniculata* (Acanthaceae), *Genlisea aurea* (Lentibulariaceae), *Actinidia chinensis* (Actinidiaceae), *Arabidopsis thaliana* (Brassicaceae), *Cucumis sativus* (Cucurbitaceae), *Glycine max* (Leguminosae), *Ipomoea nil* (Convolvulaceae), *Vitis vinifera* (Vitaceae) and *Solanum lycopersicum* (Solanaceae). To cluster protein-coding genes into families, protein sequences from the longest transcripts of each gene from *H. difformis* and other species were extracted and aligned to each other using BLAST program with a maximal e-value of $1e^{-5}$. The genome sequences and corresponding protein sequences were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/). Genes with identity< 30%, coverage< 50% were filtered out to exclude putative fragmented genes. The OrthoMCL (v14-137) (Li et al.,

2003) method was used to cluster genes from these different species into gene families. To reveal phylogenetic relationships among *H. difformis* and other closely related species, protein sequences from 196 single-copy orthologue genes were used for phylogenetic tree reconstruction. The protein sequences of the single-copy orthologue genes were aligned with MUSCLE (v3.8.31) program (Edgar, 2004), and the corresponding coding sequence alignments were generated and concatenated with the guidance of protein alignment. RAxML (v8.2.11) (Stamatakis, 2014) and TimeTree (Hedges et al., 2006) were used to construct the phylogenetic tree with the maximum likelihood (ML) method.

Gene family expansion and contraction analysis

Based on the identified gene families and the constructed phylogenetic tree with the predicted divergence time of those species, we used CAFÉ (Han et al., 2013) to analyze gene family expansion and contraction. A random birth or death model is proposed to study gene gain or loss in gene families across a specified phylogenetic tree. Then, a conditional *P*-value was calculated for each gene family. These expanded and contracted gene families in *H. difformis* ($P < 0.05$) were mapped to GO or KEGG pathways for functional enrichment analysis.

Transcription factors prediction and phylogenetic analysis

Protein sequences of *Azolla filiculoides* were downloaded from Fernbase (https://www.fernbase.org/). Protein sequences of *Nymphaea colorata*, *A. paniculata*, *Ceratophyllum demersum*, *Aldrovanda vesiculosa*, *Erythranthe guttata*, *Boea hygrometrica*, *G. aurea* and *S. indicum* were downloaded from NCBI (https://www.ncbi.nlm.nih.gov/) and Phytozome (https://phytozome.jgi.doe.gov/). These protein sequences were then analyzed by the online software PlantTFDB (http://planttfdb.gao-lab.org/), and TFs of *Spirodela polyrhiza*, *Zostera marina*, *Nelumbo nucifera* and *Utricularia gibba* were downloaded from the data source of PlantTFDB. The phylogenetic tree of different TCP proteins was constructed in MEGA11 (http://www.megasoftware.net/) and displayed using iTOL (https://itol.embl.de/).

Collinearity and WGD analysis

BLASTp (e-value <$1e^{-5}$) was used to detect orthologous genes in *H. difformis* and *A. paniculata*, to assess the level of synteny between inter-species. Then, syntenic paralogous blocks were identified with MCScanX (Wang et al., 2012) and visualized by TBtools (Chen et al., 2020). MCScanX analysis was then performed with *H. difformis* and *A. paniculata* for *Ks* evaluation. We extracted all the paralogous and orthologous gene pairs from syntenic blocks in those species to further estimate the *Ks* distribution and infer WGD dates by assuming a molecular clock (Schmutz et al., 2010).

RNA sequencing, gene expression analysis and co-expression analyses

Samples of plant shoots grown under different treatment were collected for RNA-seq, and each treatment had at least three plants. We used SOAPnuke (Chen et al., 2017) to filter the raw reads, and used HISAT2 (Kim et al., 2015) to align these filtered clean reads to our chromosome-scale genome. Read counts of these samples were calculated, and differential gene expression analysis was conducted using the R package DESeq2 (Love et al., 2014) (false discovery rate [FDR] < 0.05 and expression-level $\log_2$ [fold-change] > 1). To identify relationships between DEGs, co-expression networks were constructed using the WGCNA package (Langfelder and Horvath, 2008), and Cytoscape (Shannon et al., 2003) was used to display the network. Expression heatmaps were generated using TBtools software (Chen et al., 2020).

Quantification of phytohormones and NPA treatments

To measure phytohormone levels, the plants were grown in growth chambers at 23°C under a white light flux density of 60 μmol m^{-2} s^{-1} for 1 month. Shoots including leaf primordia were dissected from the plants and frozen in liquid nitrogen immediately after sampling. Extraction and quantification of endogenous hormones (IAA and ABA) were performed as described (Nguyen et al., 2018). The experiments were conducted in triplicate from three independent plants. For naphthylphthalamic acid (NPA) treatments, plants grown in submerged conditions were shifted to an aquarium, which contains a 10 μM NPA solution. After 1-month of treatment, emerging leaves were harvested for morphological analysis.

ROS detection

Terrestrial or submerged leaves were harvested and stained with DAB and NBT solution for H_2O_2 and O_2^- analysis as previously described (Chen et al., 2021).

Gene functional studies and expression analysis

To amplify the cDNA sequence of *HdLMI1A*, *HdLMI1B* and *HdLMI1C*, the total RNA was extracted from shoots including leaf primordia of plants grown for a month and used to synthesize cDNA, same as described previously (Li et al., 2020). Amplified fragments of expected length were purified and cloned into a pEASY-T1 cloning vector (TransGen, China) for sequencing. The function of *HdLMI1A*, *HdLMI1B* and *HdLMI1C* were investigated through the VIGS technique: Specific fragments (171 bp for *HdLMI1A*, 268 bp for *HdLMI1B* and 252 bp for *HdLMI1C*) were amplified, purified, and independently introduced into the pTRV2 vector (http://www.bt-lab.cn/) by Trelief SoSoo Cloning Kit (Tsingke Biological Technology, China) to generate constructs TRV2-HdLMI1A, TRV2-HdLMI1B and TRV2-HdLMI1C; Recombinant vectors were then introduced into *Agrobacterium tumefaciens* LBA4404 strain by the liquid nitrogen freeze-thaw method; At least twenty wild-type individuals for each treatment. After two rounds of infiltration (7 days and 14 days), plants were kept growing for 3 days in terrestrial conditions and then shifted to aquariums. After 20 days of growth in submerged conditions, representative leaves from ten individuals showing phenotypic changes were photo-documented for analysis. The silencing efficiency of VIGS treatments was determined by quantitative RT-PCR (qRT-PCR) as previously described (Li et al., 2022). The cDNAs from individual developing leaves of the VIGS-treated plants were used as templates. The primers used in this study are provided in Table S13. The significant difference in the relative expression levels between the mock and the transgenic lines were evaluated using Dunnett's test with the two-sided function in using IBM SPSS version 24 to investigate the significant differences among samples at $P < 0.01$.

Acknowledgements

We thank professor Shilin Chen of the Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences for kindly providing the genome

annotation of *A. paniculata* and Dr. Markus Stetter, University of Cologne, Germany for critical reading of the manuscript. This study was supported by the National Key Research and Development Program of China (2017YFE0128800), the National Natural Science Foundation of China (31870384, 32101254), and the International Partnership Program of the Chinese Academy of Sciences (152342KYSB20200021).

Author contributions

G.J.L. and H.W.H. designed the experiments and wrote the manuscript; G.J.L., X.Y.Z., J.J.Y., Y.M.C., and Y.C. performed genome and transcriptome analyses; Y.X.L., Z.L.S., and M.L.X. assembled the figures; X.Y.W., K.S.L., and X.Z.L. conducted environmental treatment experiments and collected the materials; S.Q.H., J.P., S.K., K.U.T., and I.H. improved the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare that they have no competing interest.

Data availability

The genome dataset is available at the GenBank with accession number PRJNA872356. Genome sequencing data (BGI seq reads, PacBio Sequel reads and Hi-C interaction reads) and environmental treatment of RNA-Seq reads are available in the NCBI Sequence Read Archive under accessions PRJNA855313 and PRJNA859677.

References

Agarie, S., Kai, M., Takatsuji, H. and Ueno, O. (1997) Expression of C3 and C4 photosynthetic characteristics in the amphibious plant Eleocharis vivipara: structure and analysis of the expression of isogenes for pyruvate, orthophosphate dikinase. *Plant Mol Biol* **34**, 363-369.

Akman, M., Bhikharie, A.V., McLean, E.H., Boonman, A., Visser, E.J., Schranz, M.E. and van Tienderen, P.H. (2012) Wait or escape? Contrasting submergence tolerance strategies of Rorippa amphibia, Rorippa sylvestris and their hybrid. *Ann Bot* **109**, 1263-1276.

Albert, V.A. and Renner, T. (2020) Aquatic angiosperm ambiguities answered. *Nature Plants* **6**, 181-183.

Amano, R., Nakayama, H., Momoi, R., Omata, E., Gunji, S., Takebayashi, Y., Kojima, M., Ikematsu, S., Ikeuchi, M., Iwase, A., Sakamoto, T., Kasahara, H., Sakakibara, H., Ferjani, A. and Kimura, S. (2020) Molecular Basis for Natural

Vegetative Propagation via Regeneration in North American Lake Cress, Rorippa aquatica (Brassicaceae). *Plant Cell Physiol* **61**, 353-369.

Andres, R.J., Coneva, V., Frank, M.H., Tuttle, J.R., Samayoa, L.F., Han, S.W., Kaur, B., Zhu, L.L., Fang, H., Bowman, D.T., Rojas-Pierce, M., Haigler, C.H., Jones, D.C., Holland, J.B., Chitwood, D.H. and Kuraparthy, V. (2017) Modifications to a LATE MERISTEM IDENTITY1 gene are responsible for the major leaf shapes of Upland cotton (Gossypium hirsutum L.). *Proceedings of the National Academy of Sciences of the United States of America* **114**, E57-E66.

Bailey-Serres, J., Lee, S.C. and Brinton, E. (2012) Waterproofing crops: effective flooding survival strategies. *Plant Physiol* **160**, 1698-1709.

Bailey-Serres, J. and Voesenek, L.A. (2008) Flooding stress: acclimations and genetic diversity. *Annu Rev Plant Biol* **59**, 313-339.

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.C., Estreicher, A., Gasteiger, E., Martin, M.J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S. and Schneider, M. (2003) The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res* **31**, 365-370.

Botts, P.S., Lawrence, J.M., Witz, B.W. and Kovach, C.W. (1990) Plasticity in morphology, proximate composition, and energy content of Hygrophila-polysperma (Roxb) Anders. *Aquat Bot* **36**, 207-214.

Burton, J.N., Adey, A., Patwardhan, R.P., Qiu, R., Kitzman, J.O. and Shendure, J. (2013) Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol* **31**, 1119-1125.

Cantarel, B.L., Korf, I., Robb, S.M.C., Parra, G., Ross, E., Moore, B., Holt, C., Alvarado, A.S. and Yandell, M. (2008) MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research* **18**, 188-196.

Cenci, A. and Rouard, M. (2017) Evolutionary analyses of GRAS transcription factors in angiosperms. *Front Plant Sci* **8**, 273.

Chen, C.J., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y.H. and Xia, R. (2020) TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant* **13**, 1194-1202.

Chen, Y., Li, G.J., Yang, J.J., Zhao, X.Y., Sun, Z.L. and Hou, H.W. (2021) Role of Nramp transporter genes of Spirodela polyrhiza in cadmium accumulation. *Ecotox Environ Safe* **227**.

Chen, Y.X., Chen, Y.S., Shi, C.M., Huang, Z.B., Zhang, Y., Li, S.K., Li, Y., Ye, J., Yu, C., Li, Z., Zhang, X.Q., Wang, J., Yang, H.M., Fang, L. and Chen, Q. (2017) SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience* **7**.

Cheng, J.L., Li, J., Zhang, Z., Lu, H., Chen, G.Q., Yao, B.B., Dong, Y.X., Ma, L., Yuan, X.X., Xu, J.X., Zhang, Y., Dai, W.M., Yang, X.H., Xue, L.F., Zhang, Y., Zhang, C.B., Mauricio, R., Peng, G., Hu, S.J., Valverde, B.E., Song, X.L., Li, Y., Stift, M. and Qiang, S. (2021a) Autopolyploidy-driven range expansion of a temperate-originated plant to pan-tropic under global change. *Ecol Monogr*

91.
Cheng, Z.Y., Luan, Y.T., Meng, J.S., Sun, J., Tao, J. and Zhao, D.Q. (2021b) WRKY transcription factor response to high-temperature stress. *Plants-Basel* **10**.
Debbarma, J., Sarki, Y.N., Saikia, B., Boruah, H.P.D., Singha, D.L. and Chikkaputtaiah, C. (2019) Ethylene Response Factor (ERF) family proteins in abiotic stresses and CRISPR-Cas9 genome editing of ERFs for multiple abiotic stress tolerance in crop plants: a review. *Mol Biotechnol* **61**, 153-172.
Debernardi, J.M., Mecchia, M.A., Vercruyssen, L., Smaczniak, C., Kaufmann, K., Inze, D., Rodriguez, R.E. and Palatnik, J.F. (2014) Post-transcriptional control of GRF transcription factors by microRNA miR396 and GIF co-activator affects leaf size and longevity. *Plant J* **79**, 413-426.
Deng, X., Yang, D., Sun, H., Liu, J., Song, H., Xiong, Y., Wang, Y., Ma, J., Zhang, M., Li, J., Liu, Y. and Yang, M. (2022) Time-course analysis and transcriptomic identification of key response strategies to complete submergence in Nelumbo nucifera. *Hortic Res*.
Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S. and Aiden, E.L. (2016) Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst* **3**, 95-98.
Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.
Feller, A., Machemer, K., Braun, E.L. and Grotewold, E. (2011) Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant Journal* **66**, 94-116.
Han, M.V., Thomas, G.W.C., Lugo-Martinez, J. and Hahn, M.W. (2013) Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol Biol Evol* **30**, 1987-1997.
Hareven, D., Gutfinger, T., Parnis, A., Eshed, Y. and Lifschitz, E. (1996) The making of a compound leaf: genetic manipulation of leaf architecture in tomato. *Cell* **84**, 735-744.
He, D., Guo, P., Gugger, P.F., Guo, Y., Liu, X. and Chen, J. (2018) Investigating the molecular basis for heterophylly in the aquatic plant Potamogeton octandrus (Potamogetonaceae) with comparative transcriptomics. *PeerJ* **6**, e4448.
Hedges, S.B., Dudley, J. and Kumar, S. (2006) TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* **22**, 2971-2972.
Horiguchi, G., Matsumoto, K., Nemoto, K., Inokuchi, M. and Hirotsu, N. (2021) Transition from proto-kranz-type photosynthesis to HCO3- use photosynthesis in the amphibious plant Hygrophila polysperma. *Frontiers in Plant Science* **12**.
Huang, X.N., Li, S.G., Gao, Y.C. and Zhan, A.B. (2018) Genome-wide identification, characterization and expression analyses of Heat Shock Protein-related genes in a highly invasive ascidian Ciona savignyi. *Front Physiol* **9**.
Hunt, L., Bailey, K.J. and Gray, J.E. (2010) The signalling peptide EPFL9 is a positive regulator of stomatal development. *New Phytologist* **186**, 609-614.
Ingale, K.G., Thakurdesai, P.A. and Vyawahare, N.S. (2013) Protective effect of Hygrophila spinosa against cisplatin induced nephrotoxicity in rats. *Indian J*

Pharmacol **45**, 232-236.
Jha, P., Ochatt, S.J. and Kumar, V. (2020) WUSCHEL: a master regulator in plant growth signaling. *Plant Cell Rep* **39**, 431-444.
Jin, J., Tian, F., Yang, D.C., Meng, Y.Q., Kong, L., Luo, J. and Gao, G. (2017) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res* **45**, D1040-D1045.
Jung, J.H., Barbosa, A.D., Hutin, S., Kumita, J.R., Gao, M., Derwort, D., Silva, C.S., Lai, X., Pierre, E., Geng, F., Kim, S.B., Baek, S., Zubieta, C., Jaeger, K.E. and Wigge, P.A. (2020) A prion-like domain in ELF3 functions as a thermosensor in Arabidopsis. *Nature* **585**, 256-260.
Karatas, M., Aasim, M., Cinar, A. and Dogan, M. (2013) Adventitious shoot regeneration from leaf explant of dwarf Hygro (Hygrophila polysperma (Roxb.) T. Anderson). *Sci World J*.
Kieffer, M., Master, V., Waites, R. and Davies, B. (2011) TCP14 and TCP15 affect internode length and leaf shape in Arabidopsis. *Plant J* **68**, 147-158.
Kim, D., Langmead, B. and Salzberg, S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* **12**, 357-360.
Kim, J., Joo, Y., Kyung, J., Jeon, M., Park, J.Y., Lee, H.G., Chung, D.S., Lee, E. and Lee, I. (2018) A molecular basis behind heterophylly in an amphibious plant, Ranunculus trichophyllus. *Plos Genetics* **14**, e1007208.
Kimura, S., Koenig, D., Kang, J., Yoong, F.Y. and Sinha, N. (2008) Natural variation in leaf morphology results from mutation of a novel KNOX gene. *Curr Biol* **18**, 672-677.
Koga, H., Kojima, M., Takebayashi, Y., Sakakibara, H. and Tsukaya, H. (2021) Identification of the unique molecular framework of heterophylly in the amphibious plant Callitriche palustris L. *Plant Cell* **33**, 3272-3292.
Kordyum, E.L. and Mosyakin, S.L. (2020) Endosperm of angiosperms and genomic imprinting. *Life (Basel)* **10**.
Koyama, T., Mitsuda, N., Seki, M., Shinozaki, K. and Ohme-Takagi, M. (2010) TCP transcription factors regulate the activities of ASYMMETRIC LEAVES1 and miR164, as well as the auxin response, during differentiation of leaves in Arabidopsis. *Plant Cell* **22**, 3574-3588.
Kumar, R., Kushalappa, K., Godt, D., Pidkowich, M.S., Pastorelli, S., Hepworth, S.R. and Haughn, G.W. (2007) The Arabidopsis BEL1-LIKE HOMEODOMAIN proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. *Plant Cell* **19**, 2719-2735.
Kumar, S., Stecher, G., Suleski, M. and Hedges, S.B. (2017) TimeTree: A resource for timelines, timetrees, and divergence times. *Mol Biol Evol* **34**, 1812-1819.
Kuroha, T., Nagai, K., Gamuyao, R., Wang, D.R., Furuta, T., Nakamori, M., Kitaoka, T., Adachi, K., Minami, A., Mori, Y., Mashiguchi, K., Seto, Y., Yamaguchi, S., Kojima, M., Sakakibara, H., Wu, J., Ebana, K., Mitsuda, N., Ohme-Takagi, M., Yanagisawa, S., Yamasaki, M., Yokoyama, R., Nishitani, K., Mochizuki, T., Tamiya, G., McCouch, S.R. and Ashikari, M. (2018) Ethylene-gibberellin signaling underlies adaptation of rice to periodic flooding. *Science* **361**,

181-186.
Kuwabara, A., Ikegami, K., Koshiba, T. and Nagata, T. (2003) Effects of ethylene and abscisic acid upon heterophylly in Ludwigia arcuata (Onagraceae). *Planta* **217**, 880-887.
Lai, W., Zhu, C., Hu, Z., Liu, S., Wu, H. and Zhou, Y. (2021) Identification and transcriptional analysis of Zinc Finger-Homeodomain (ZF-HD) family genes in Cucumber. *Biochem Genet* **59**, 884-901.
Lan, T., Renner, T., Ibarra-Laclette, E., Farr, K.M., Chang, T.H., Cervantes-Perez, S.A., Zheng, C., Sankoff, D., Tang, H., Purbojati, R.W., Putra, A., Drautz-Moses, D.I., Schuster, S.C., Herrera-Estrella, L. and Albert, V.A. (2017) Long-read sequencing uncovers the adaptive topography of a carnivorous plant genome. *Proc Natl Acad Sci U S A* **114**, E4435-E4441.
Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559.
Lehmann, J., Jorgensen, M.E., Fratz, S., Muller, H.M., Kusch, J., Scherzer, S., Navarro-Retamal, C., Mayer, D., Bohm, J., Konrad, K.R., Terpitz, U., Dreyer, I., Mueller, T.D., Sauer, M., Hedrich, R., Geiger, D. and Maierhofer, T. (2021) Acidosis-induced activation of anion channel SLAH3 in the flooding-related stress response of Arabidopsis. *Curr Biol* **31**, 3575-3585 e3579.
Lehti-Shiu, M.D., Panchy, N., Wang, P., Uygun, S. and Shiu, S.H. (2017) Diversity, expansion, and evolutionary novelty of plant DNA-binding transcription factor families. *Biochim Biophys Acta Gene Regul Mech* **1860**, 3-20.
Li, G., Hu, S., Hou, H. and Kimura, S. (2019) Heterophylly: Phenotypic plasticity of leaf shape in aquatic and amphibious plants. *Plants (Basel)* **8**, 420.
Li, G., Hu, S., Yang, J., Schultz, E.A., Clarke, K. and Hou, H. (2017) Water-Wisteria as an ideal plant to study heterophylly in higher aquatic plants. *Plant Cell Reports* **36**, 1225-1236.
Li, G., Hu, S., Yang, J., Zhao, X., Kimura, S., Schultz, E.A. and Hou, H. (2020) 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.
Li, G., Hu, S., Zhao, X., Kumar, S., Li, Y., Yang, J. and Hou, H. (2021) Mechanisms of the morphological plasticity induced by phytohormones and the environment in plants. *Int J Mol Sci* **22**, 765.
Li, G., Yang, J., Chen, Y., Zhao, X., Chen, Y., Kimura, S., Hu, S. and Hou, H. (2022) SHOOT MERISTEMLESS participates in the heterophylly of Hygrophila difformis (Acanthaceae). *Plant Physiol*.
Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595.
Li, L., Stoeckert, C.J. and Roos, D.S. (2003) OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research* **13**, 2178-2189.
Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550.
Lu, R.S., Chen, Y., Zhang, X.Y., Feng, Y., Comes, H.P., Li, Z., Zheng, Z.S., Yuan, Y.,

Wang, L.Y., Huang, Z.J., Guo, Y., Sun, G.P., Olsen, K.M., Chen, J. and Qiu, Y.X. (2021) Genome sequencing and transcriptome analyses provide insights into the origin and domestication of water caltrop (Trapa spp., Lythraceae). *Plant Biotechnol J*.

Majoros, W.H., Pertea, M. and Salzberg, S.L. (2004) TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* **20**, 2878-2879.

Mao, H.D., Wang, H.W., Liu, S.X., Li, Z., Yang, X.H., Yan, J.B., Li, J.S., Tran, L.S.P. and Qin, F. (2015) A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. *Nature Communications* **6**.

Mathew, I.E. and Agarwal, P. (2018) May the fittest protein evolve: favoring the plant-specific origin and expansion of NAC transcription factors. *Bioessays* **40**.

Mistry, J., Bateman, A. and Finn, R.D. (2007) Predicting active site residue annotations in the Pfam database. *Bmc Bioinformatics* **8**.

Mitchell, A., Chang, H.Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S., Sangrador-Vegas, A., Scheremetjew, M., Rato, C., Yong, S.Y., Bateman, A., Punta, M., Attwood, T.K., Sigrist, C.J.A., Redaschi, N., Rivoire, C., Xenarios, I., Kahn, D., Guyot, D., Bork, P., Letunic, I., Gough, J., Oates, M., Haft, D., Huang, H.Z., Natale, D.A., Wu, C.H., Orengo, C., Sillitoe, I., Mi, H.Y., Thomas, P.D. and Finn, R.D. (2015) The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res* **43**, D213-D221.

Momokawa, N., Kadono, Y. and Kudoh, H. (2011) Effects of light quality on leaf morphogenesis of a heterophyllous amphibious plant, Rotala hippuris. *Ann Bot* **108**, 1299-1306.

Muller, J.T., van Veen, H., Bartylla, M.M., Akman, M., Pedersen, O., Sun, P., Schuurink, R.C., Takeuchi, J., Todoroki, Y., Weig, A.R., Sasidharan, R. and Mustroph, A. (2021) Keeping the shoot above water - submergence triggers antithetical growth responses in stems and petioles of watercress (Nasturtium officinale). *New Phytol* **229**, 140-155.

Nagai, K., Mori, Y., Ishikawa, S., Furuta, T., Gamuyao, R., Niimi, Y., Hobo, T., Fukuda, M., Kojima, M., Takebayashi, Y., Fukushima, A., Himuro, Y., Kobayashi, M., Ackley, W., Hisano, H., Sato, K., Yoshida, A., Wu, J., Sakakibara, H., Sato, Y., Tsuji, H., Akagi, T. and Ashikari, M. (2020) Antagonistic regulation of the gibberellic acid response during stem growth in rice. *Nature* **584**, 109-114.

Nakayama, H. and Kimura, S. (2015) Leaves may function as temperature sensors in the heterophylly of Rorippa aquatica (Brassicaceae). *Plant Signaling & Behavior* **10**.

Nakayama, H., Nakayama, N., Seiki, S., Kojima, M., Sakakibara, H., Sinha, N. and Kimura, S. (2014) Regulation of the KNOX-GA gene module induces heterophyllic alteration in North American lake cress. *Plant Cell* **26**, 4733-4748.

Nakayama, H., Sinha, N.R. and Kimura, S. (2017) How do plants and phytohormones accomplish heterophylly, leaf phenotypic plasticity, in response to environmental cues. *Frontiers in Plant Science* **8**, 1717.

Nguyen, T.N., Tuan, P.A., Mukherjee, S., Son, S. and Ayele, B.T. (2018) Hormonal regulation in adventitious roots and during their emergence under waterlogged conditions in wheat. *Journal of Experimental Botany* **69**, 4065-4082.

Park, H., Park, J.H., Lee, Y., Woo, D.U., Jeon, H.H., Sung, Y.W., Shim, S., Kim, S.H., Lee, K.O., Kim, J.Y., Kim, C.K., Bhattacharya, D., Yoon, H.S. and Kang, Y.J. (2021) Genome of the world's smallest flowering plant, Wolffia australiana, helps explain its specialized physiology and unique morphology. *Communications Biology* **4**.

Prancl, J., Kaplan, Z., Travnicek, P. and Jarolimova, V. (2014) Genome size as a key to evolutionary complex aquatic plants: polyploidy and hybridization in Callitriche (Plantaginaceae). *Plos One* **9**.

Raschke, A., Ibanez, C., Ullrich, K.K., Anwer, M.U., Becker, S., Glockner, A., Trenner, J., Denk, K., Saal, B., Sun, X., Ni, M., Davis, S.J., Delker, C. and Quint, M. (2015) Natural variants of ELF3 affect thermomorphogenesis by transcriptionally modulating PIF4-dependent auxin response genes. *BMC Plant Biol* **15**, 197.

Ren, G., Jiang, Y., Li, A., Yin, M., Li, M., Mu, W., Wu, Y. and Liu, J. (2022) The genome sequence provides insights into salt tolerance of Achnatherum splendens (Gramineae), a constructive species of alkaline grassland. *Plant Biotechnol J* **20**, 116-128.

Sakai, H., Honma, T., Aoyama, T., Sato, S., Kato, T., Tabata, S. and Oka, A. (2001) ARR1, a transcription factor for genes immediately responsive to cytokinins. *Science* **294**, 1519-1521.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J.X., Mitros, T., Nelson, W., Hyten, D.L., Song, Q.J., Thelen, J.J., Cheng, J.L., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T., Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S.Q., Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J.C., Tian, Z.X., Zhu, L.C., Gill, N., Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing, R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C. and Jackson, S.A. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**, 178-183.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504.

Sharma, V., Gupta, R.C., Bala, S. and Singh, B. (2013) New chromosome number reports in some medicinally important angiosperms of North India. *Cytologia* **78**, 285-296.

Shi, T., Huneau, C., Zhang, Y., Li, Y., Chen, J., Salse, J. and Wang, Q. (2022) The slow-evolving Acorus tatarinowii genome sheds light on ancestral monocot

evolution. *Nat Plants* **8**, 764-777.

Simao, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V. and Zdobnov, E.M. (2015) BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212.

Slater, G.S. and Birney, E. (2005) Automated generation of heuristics for biological sequence comparison. *Bmc Bioinformatics* **6**.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312-1313.

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S. and Morgenstern, B. (2006) AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res* **34**, W435-W439.

Sun, W., Leng, L. and Yin, Q. (2019) The genome of the medicinal plant Andrographis paniculata provides insight into the biosynthesis of the bioactive diterpenoid neoandrographolide (vol 97, pg 841, 2018). *Plant Journal* **97**, 996-996.

Swaminathan, K., Peterson, K. and Jack, T. (2008) The plant B3 superfamily. *Trends Plant Sci* **13**, 647-655.

Tran, D., Girault, T., Guichard, M., Thomine, S., Leblanc-Fournier, N., Moulia, B., de Langre, E., Allain, J.M. and Frachisse, J.M. (2021) Cellular transduction of mechanical oscillations in plants by the plasma-membrane mechanosensitive channel MSL10. *Proc Natl Acad Sci U S A* **118**.

van Veen, H., Mustroph, A., Barding, G.A., Vergeer-van Eijk, M., Welschen-Evertman, R.A., Pedersen, O., Visser, E.J., Larive, C.K., Pierik, R., Bailey-Serres, J., Voesenek, L.A. and Sasidharan, R. (2013) Two Rumex species from contrasting hydrological niches regulate flooding tolerance through distinct mechanisms. *Plant Cell* **25**, 4691-4707.

van Veen, H. and Sasidharan, R. (2021) Shape shifting by amphibious plants in dynamic hydrological niches. *New Phytol* **229**, 79-84.

Vlad, D., Kierzkowski, D., Rast, M.I., Vuolo, F., Dello Ioio, R., Galinha, C., Gan, X.C., Hajheidari, M., Hay, A., Smith, R.S., Huijser, P., Bailey, C.D. and Tsiantis, M. (2014) Leaf shape evolution through duplication, regulatory diversification, and loss of a homeobox gene. *Science* **343**, 780-783.

Wan, T., Liu, Z., Leitch, I.J., Xin, H., Maggs-Kolling, G., Gong, Y., Li, Z., Marais, E., Liao, Y., Dai, C., Liu, F., Wu, Q., Song, C., Zhou, Y., Huang, W., Jiang, K., Wang, Q., Yang, Y., Zhong, Z., Yang, M., Yan, X., Hu, G., Hou, C., Su, Y., Feng, S., Yang, J., Yan, J., Chu, J., Chen, F., Ran, J., Wang, X., Van de Peer, Y., Leitch, A.R. and Wang, Q. (2021) The Welwitschia genome reveals a unique biology underpinning extreme longevity in deserts. *Nat Commun* **12**, 4247.

Wang, H., Niu, H.H., Li, C., Shen, G.Y., Liu, X.F., Weng, Y.Q., Wu, T. and Li, Z. (2020) WUSCHEL-related homeobox1 (WOX1) regulates vein patterning and leaf size in Cucumis sativus. *Horticulture Research* **7**.

Wang, X., Zhang, J.J., Xie, Y.Y., Liu, X., Wen, L.Z., Wang, H.F., Zhang, J., Li, J., Han, L., Yu, X.L., Mysore, K.S., Wen, J.Q. and Zhou, C. (2021) LATE MERISTEM IDENTITY1 regulates leaf margin development via the auxin transporter gene

SMOOTH LEAF MARGIN1. *Plant Physiology* **187**, 218-235.

Wang, Y.P., Tang, H.B., DeBarry, J.D., Tan, X., Li, J.P., Wang, X.Y., Lee, T.H., Jin, H.Z., Marler, B., Guo, H., Kissinger, J.C. and Paterson, A.H. (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* **40**.

Whitewoods, C.D., Goncalves, B., Cheng, J., Cui, M.L., Kennaway, R., Lee, K., Bushell, C., Yu, M., Piao, C.L. and Coen, E. (2020) Evolution of carnivorous traps from planar leaves through simple shifts in gene expression. *Science* **367**, 91-+.

Wissler, L., Codoner, F.M., Gu, J., Reusch, T.B., Olsen, J.L., Procaccini, G. and Bornberg-Bauer, E. (2011) Back to the sea twice: identifying candidate plant genes for molecular evolution to marine life. *BMC Evol Biol* **11**, 8.

Wu, P., Zhang, L.K., Zhang, K., Yin, Y.L., Liu, A.L., Zhu, Y., Fu, Y., Sun, F.F., Zhao, S.P., Feng, K., Xu, X.W., Chen, X.H., Cheng, F. and Li, L.J. (2022) The adaptive evolution of Euryale ferox to the aquatic environment through paleo-hexaploidization. *Plant Journal* **110**, 627-645.

Xiao, C.L., Chen, Y., Xie, S.Q., Chen, K.N., Wang, Y., Han, Y., Luo, F. and Xie, Z. (2017) MECAT : fast mapping, error correction, and de novo assembly for single-molecule sequencing reads. *Nature Methods* **14**, 1072-+.

Yaneff, A., Vitali, V. and Amodeo, G. (2015) PIP1 aquaporins: Intrinsic water channels or PIP2 aquaporin modulators? *Febs Letters* **589**, 3508-3515.

Yu, H., Zhang, L., Wang, W., Tian, P., Wang, W., Wang, K., Gao, Z., Liu, S., Zhang, Y., Irish, V.F. and Huang, T. (2021) TCP5 controls leaf margin development by regulating KNOX and BEL-like transcription factors in Arabidopsis. *J Exp Bot* **72**, 1809-1821.

Yu, T., Bai, Y., Liu, Z., Wang, Z., Yang, Q., Wu, T., Feng, S., Zhang, Y., Shen, S., Li, Q., Gu, L. and Song, X. (2022) Large-scale analyses of heat shock transcription factors and database construction based on whole-genome genes in horticultural and representative plants. *Hortic Res*.

Yuan, X., Wang, H., Cai, J.T., Bi, Y., Li, D.Y. and Song, F.M. (2019) Rice NAC transcription factor ONAC066 functions as a positive regulator of drought and oxidative stress response. *Bmc Plant Biology* **19**.

Zhang, L., Zhou, L., Yung, W.S., Su, W. and Huang, M. (2021) Ectopic expression of Torenia fournieri TCP8 and TCP13 alters the leaf and petal phenotypes in Arabidopsis thaliana. *Physiol Plant* **173**, 856-866.

Zhang, R., Fu, X.H., Zhao, C.Y., Cheng, J., Liao, H., Wang, P.P., Yao, X., Duan, X.S., Yuan, Y., Xu, G.X., Kramer, E.M., Shan, H.Y. and Kong, H.Z. (2020) Identification of the key regulatory genes involved in elaborate petal development and specialized character formation in Nigella damascena (Ranunculaceae). *Plant Cell* **32**, 3095-3112.

Zhou, X.J., Li, J.T., Wang, H.L., Han, J.W., Zhang, K., Dong, S.W., Zhang, Y.Z., Ya, H.Y., Cheng, Y.W. and Sun, S.S. (2021) The chromosome-scale genome assembly, annotation and evolution of Rhododendron henanense subsp. lingbaoense. *Mol Ecol Resour*.

Figures

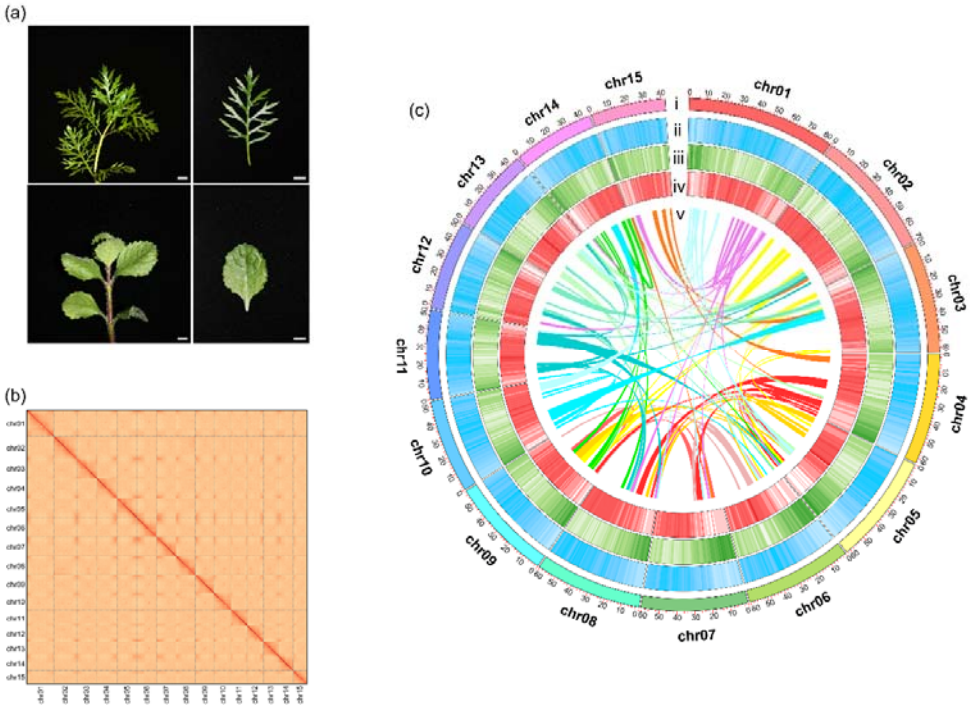

Figure 1. The plant *Hygrophila difformis* and genome assembly.

(a) Morphology of the submerged plant (up, left), submerged leaf (up, right), terrestrial plant (down, left) and terrestrial leaf (down, right) of *H. difformis*. Bars = 1 cm.

(b) High-throughput chromosome conformation capture map of the *H. difformis* genome showing strong interactive signal within intra-chromosome (intra-Chr) and at diagonal regions.

(c) Overview of *H. difformis* genome assembly. Chromosome (i), repeats (ii), gene density (iii), GC ratio (iv) and collinearity blocks (v) were displayed. The color from light to dark on the circular density scale indicates a density of features from low to high presented in a given track.

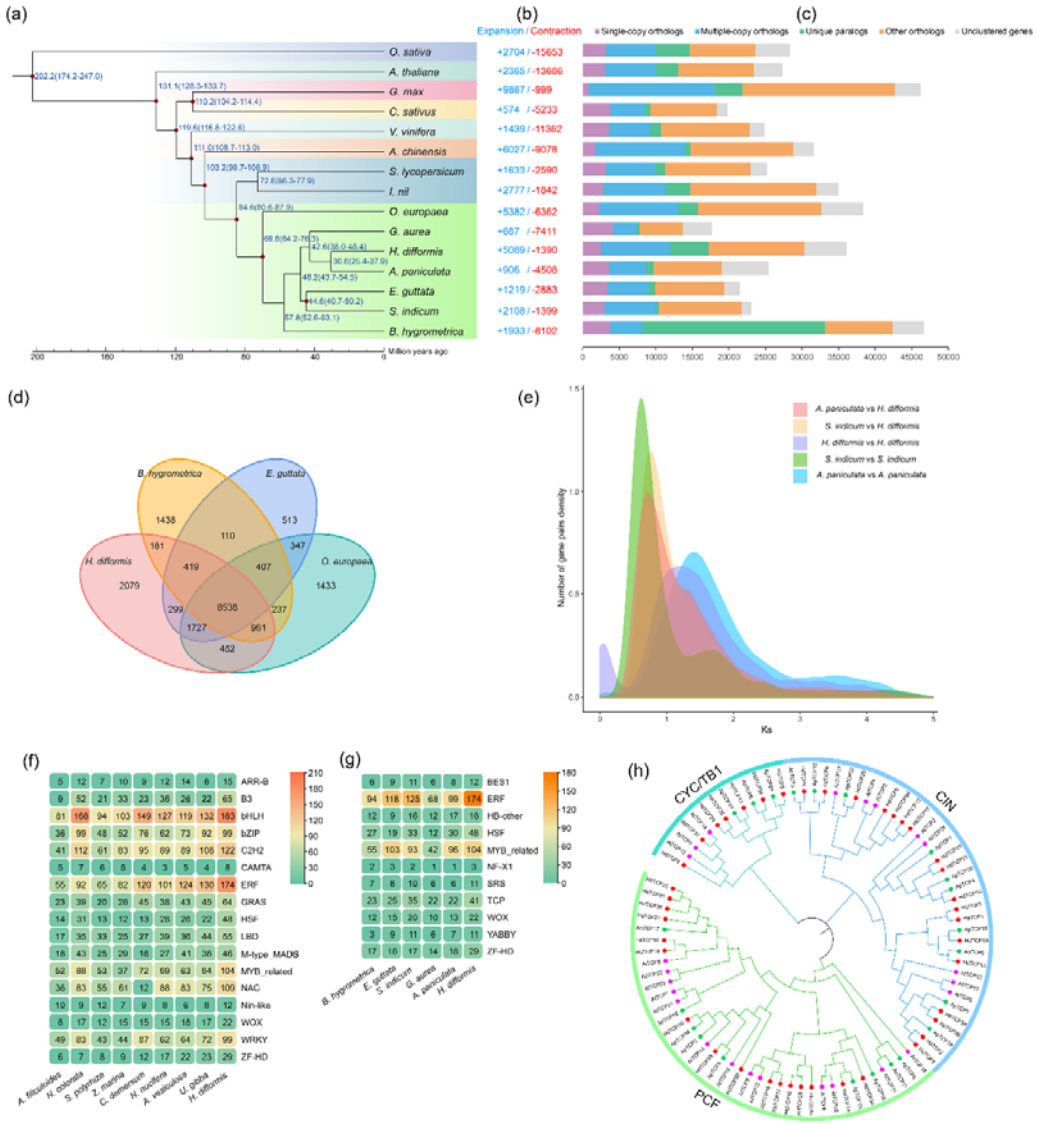

Figure 2. Evolutionary analyses of *H. difformis* genome.

(a) Phylogenetic relationships and times of divergence between *H. difformis* and other species.

(b) Expansions and contractions of gene families.

(c) Clusters of orthologous and paralogous gene families in *H. difformis* and other species. Only the longest isoform for each gene is shown.

(d) Venn diagram of orthologous genes shared among the four species *H. difformis*, *Boea hygrometrica* (Gesneriaceae), *Erythranthe guttata* (Phrymaceae), and *Olea europaea* (Oleaceae).

(e) Distribution of average synonymous substitution levels (*Ks*) between syntenic

blocks.

(f) Specific expansion of TF families in *H. difformis* compared with other plants. The heatmap shows categorized orthogroups of TF families which have expanded numbers in *H. difformis* compared with other aquatic plants (left) and land plants (right).

(g) The phylogenetic tree of TCPs among *H. difformis*, *A. paniculata*, and *A. thaliana*. The TCP proteins of *H. difformis* are marked by red solid cycle.

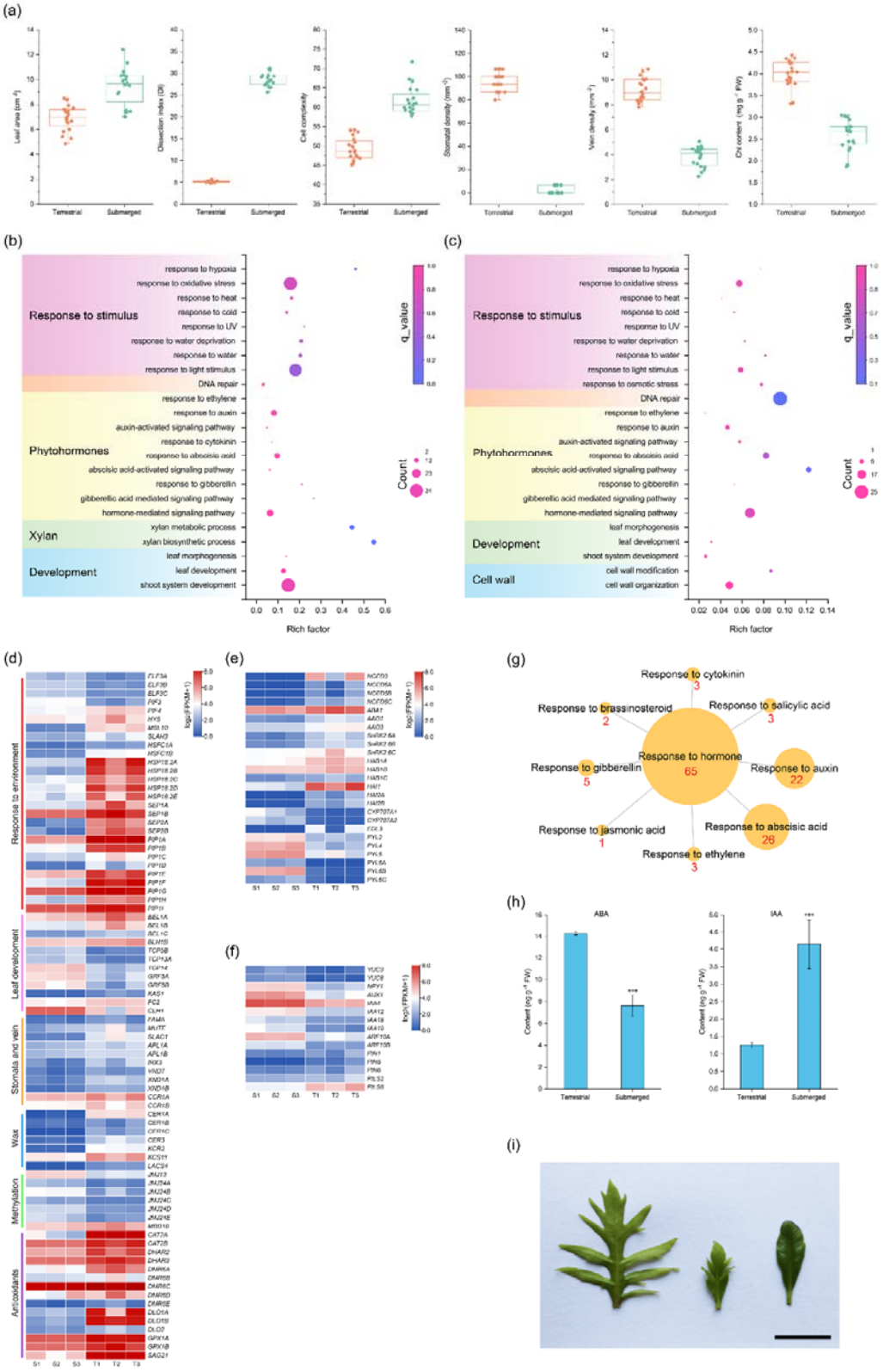

Figure 3. Phenotype and transcriptomic analyses of *H. difformis* under terrestrial or

submerged conditions.

(a) Morphological parameters (leaf area, dissection index, cell complexity, stomatal density, vein density, and chlorophyll content) of *H. difformis* under terrestrial or submerged condition. Data from twenty biological individuals each, and significant differences were detected by Student's *t*-test ($P < 0.01$).

(b) GO term enrichment of representative upregulated genes in terrestrial shoots compared to submerged shoots.

(c) GO term enrichment of representative upregulated genes in submerged shoots compared to submerged shoots.

(d) DEGs (false discovery rate [FDR] < 0.05 and expression-level $\log_2$ [fold-change] > 1) of which involved in environmental response, leaf development, stomata and vein, wax, methylation and antioxidants.

(e) DEGs which are involved in ABA metabolism and signaling.

(f) DEGs which are involved in auxin metabolism and signaling.

(g) Diagram for phytohormone-related categories showing different expressions in terrestrial or submerged conditions. The number of genes is in red, represented by the relative size of circles that belong to each gene category.

(h) ABA and IAA content in shoots of *H. difformis* under terrestrial or submerged conditions. Error bars represent ± SD. Asterisks indicate significant differences (Student's *t*-test: ***, $P < 0.001$). Data are means of three independent experiments.

(i) Effect of NPA on *H. difformis* heterophylly. Photos show the emerging leaves after 10 μM NPA treatment for 1 month. The oldest is on the left, and the youngest leaf is on the right. Photo is a representative of ten independent plants under NPA treatment. Bar= 1 cm.

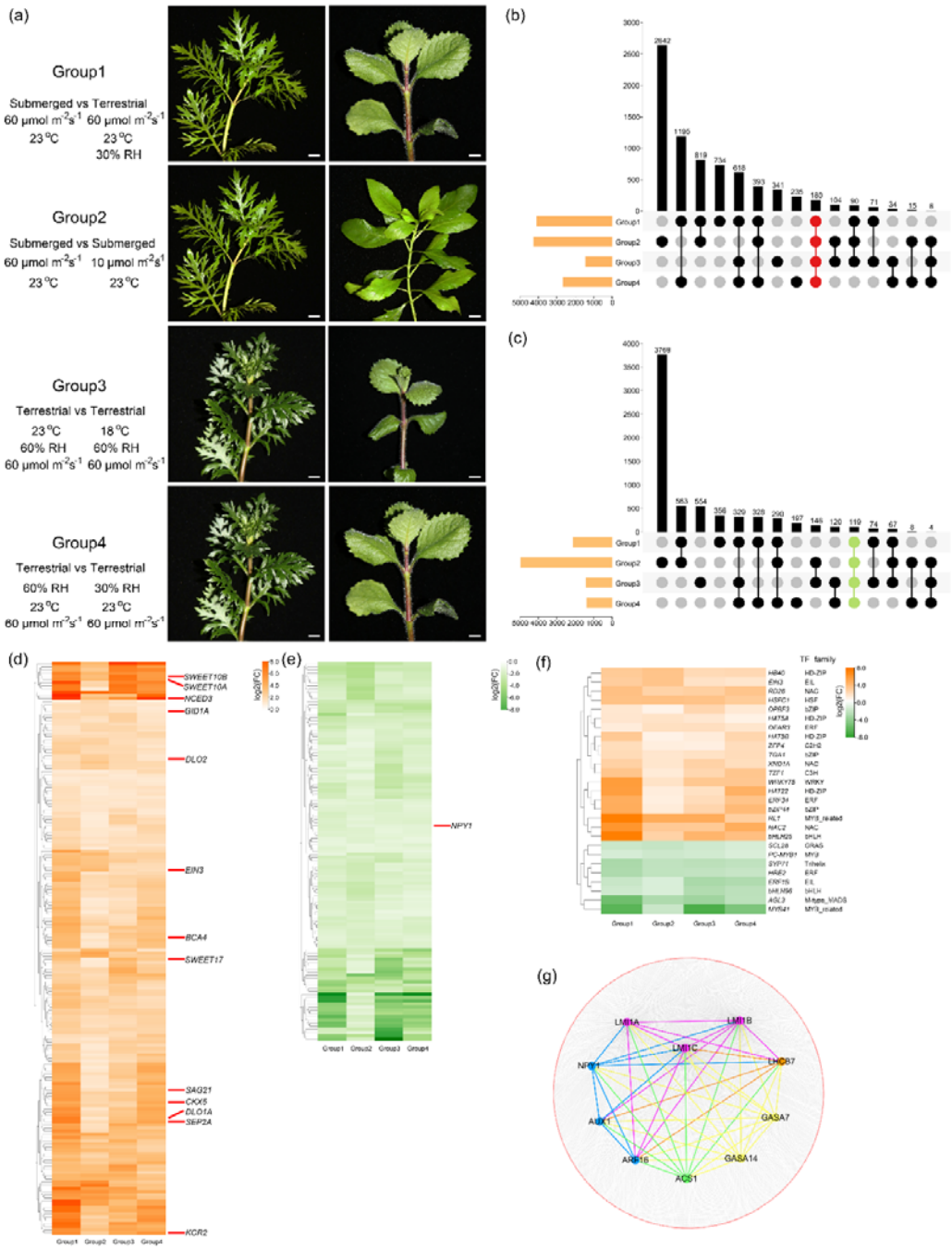

Figure 4. Phenotype and transcriptomic analyses of *H. difformis* under diverse environmental treatments.

(a) Phenotypes of *H. difformis* grown in different conditions. Group1: submerged *vs* terrestrial plants. Group2: Normal light density *vs* low light density. Group3: Normal temperature *vs* low temperature. Group4: high humidity *vs* normal humidity. Bars= 1cm.

(b-c) Comparative transcriptome analyses of *H. difformis*. UpSet plots present the overlapping DEGs in 4 comparison groups. UpSet plot of DEGs upregulated in the control (b). UpSet plot of DEGs downregulated in the control (c).

(d) Upregulated DEGs expression profiles.

(e) Downregulated DEGs expression profiles.

(f) Expression patterns of 27 putative TF genes.

(g) Co-expression network reconstruction identified *LMI1* genes clustered with the genes involved in auxin and GA signal, as well as ethylene synthesis and photosynthesis.

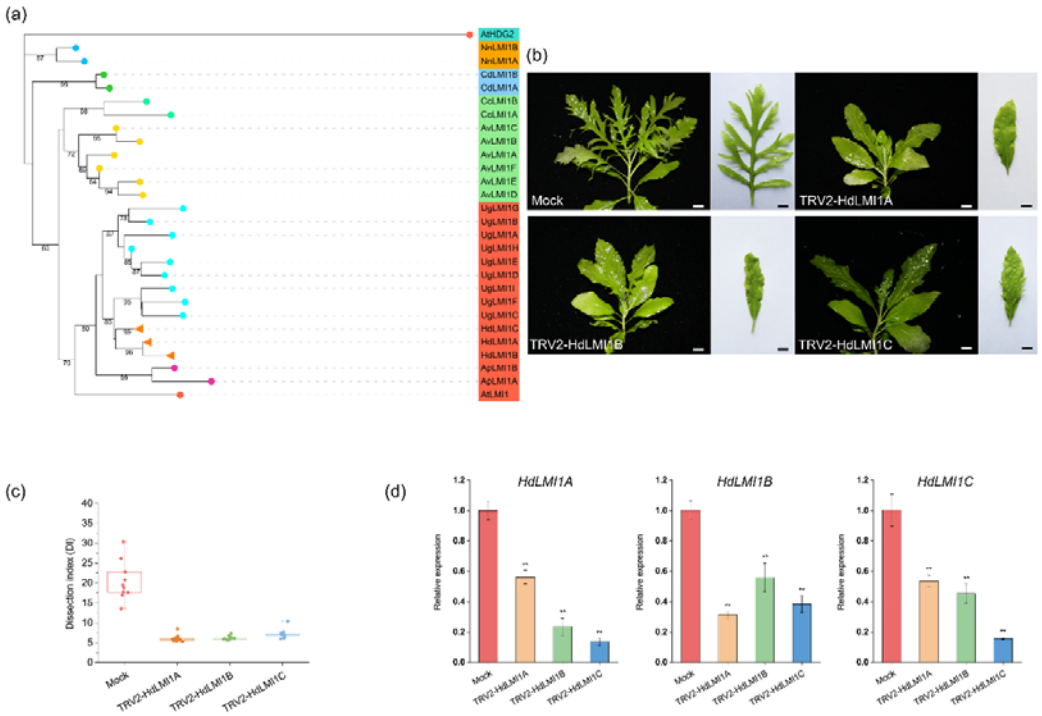

Figure 5. Identification of *LMI1* orthologs and Function of *HdLMI1s*.

(a) Phylogenetic analysis of LMI1 and its homologs. The numbers positioned at branches of the phylogenetic tree indicated bootstrap values. The AtHDG2 of *A. thaliana* was set as outroot. Triangle symbols indicate LMI1 homologs in *H. difformis*.

(b) Phenotypes of the whole plant and representative leaf in the mock, TRV2-HdLMI1A, TRV2-HdLMI1B and TRV2-HdLMI1C transgenic plants grown in submerged conditions. Bars= 1 cm.

(c) Dissection index of in the mock, TRV2-HdLMI1A, TRV2-HdLMI1B and TRV2-HdLMI1C transgenic plants grown in submerged condition. Data from ten biological individuals each, and significant differences of mock and transgenic plants were detected by Dunnett's test ($P < 0.01$).

(d) Relative expression of *HdLMI1A*, *HdLMI1B* and *HdLMI1C* in the mock, TRV2-HdLMI1A, TRV2-HdLMI1B and TRV2-HdLMI1C transgenic plants. Error bars represent ± SD. Data are means of three independent experiments. Asterisks indicate significant differences (Dunnett's test: **, $P < 0.01$).

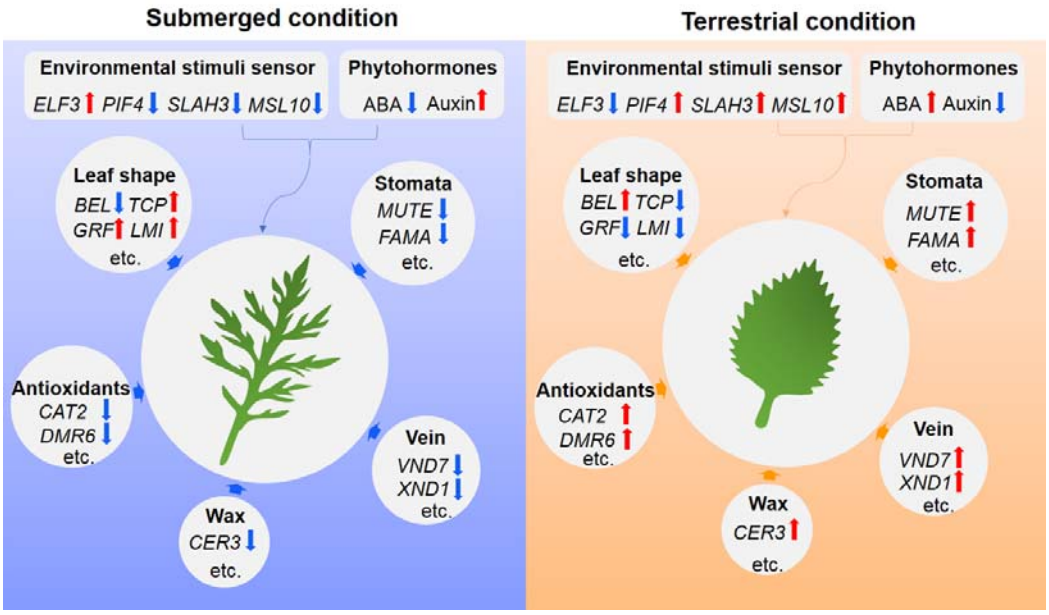

Figure 6. Schematic model of the regulation of heterophylly in *H. difformis*.

Tables

Table 1. Assembly features for the genome of *Hygrophila difformis*

Assembly feature	Value
Total length of contigs (Mb)	831.53
Contig N50 (Mb)	0.67
Scaffold N50 (Mb)	57.63
GC content	39.1%
Chromosome number	15
Total length of Chromosome (Mb)	832.72
Gene number	36,099