TITLE: *Marchantia polymorpha* gene expression atlas reveals the hierarchy of abiotic stress responses and conservation of diurnal gene expression

AUTHORS: Qiao Wen Tan[1], Peng Ken Lim[1], Zhong Chen[2], Asher Pasha[3], Nicholas Provart[3], Marek Mutwil[1*]

[1]School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore, 637551, Singapore

[2]Amoeba Education Hub, 1 West Coast Road, 128020, Singapore

[3]Department of Cell and Systems Biology/Centre for the Analysis of Genome Evolution and Function, University of Toronto, Toronto, Ontario M5S 3B2, Canada

*Corresponding author:
Marek Mutwil
School of Biological Sciences,
Nanyang Technological University,60 Nanyang Drive,
637551, Singapore,
Singapore
Email: mutwil@ntu.edu.sg

Abstract

Abiotic stress has a dramatic impact on ecosystems and the yield of crops, and global warming will likely result in more frequent and intense stresses. However, due to the complexity of the underlying signalling pathways and the fact that stresses often occur in combinations in nature, our knowledge of how plants acclimatize to abiotic stress is still lacking. Here, we used a plant with minimal regulatory network redundancy, *Marchantia polymorpha*, to study how seven abiotic stresses alone and in combination affect the phenotype, gene expression, and activity of biological pathways. Worryingly, we see no evidence of conservation of gene expression to abiotic stress between model organisms, suggesting that each species develops unique strategies to cope with abiotic stress. Interestingly, we observed that certain stresses are able to suppress others, while specific combinations can elucidate novel transcriptomic responses. Finally, we provide two online databases to study abiotic stress responses and diurnal gene expression data in Marchantia.

Introduction

The colonization of land by plants, which occurred around 470 Ma, was essential to establish habitable environments on land for all kingdoms of life [1]. Bryophytes (which include mosses, liverworts, and hornworts) represent the earliest diverging group of non-vascular land plants [2–4]. Morphology of the earliest land plant fossils (which consists primarily of tissue fragments and spores from the Middle Ordovician around 470 Ma) showed early land plants were liverwort-like [5,6]. As a liverwort, *Marchantia polymorpha* is a valuable model to better understand the emergence and evolution of land plants, as it allows us to compare the biology of aquatic algae, non-vascular plants to vascular, seed, and flowering plants. Studying the Marchantia biology can reveal the reason behind the successful terrestrialization event, as *Marchantia* contains traits essential for terrestrialization and increased complexity (e.g., hormones auxin, jasmonate, salicylic acid, protection mechanisms against desiccation, photooxidative damage)[7], albeit at a less complex organization than vascular plants [8].

Apart from its interesting evolutionary position among land plants, Marchantia is a valuable model to study basic plant biology. Largely due to the lack of whole-genome duplications in the liverwort lineage, *Marchantia* shows a simpler, redundancy-free regulatory genome [8], that together with the ease of growth and genetic manipulation [9], make Marchantia an excellent model to study general plant biology. The *Marchantia* genome contains all necessary components for most land-plant signalling pathways, but compared to flowering plants, the components are reduced to the minimum number, making it easier to dissect the pathways [8]. For example, the auxin signalling network in *Marchantia* is simple yet functional, with all relevant genes existing as single orthologs [10]. Similarly, cellulose biosynthesis in Marchantia uses the same but simplified machinery; while *Arabidopsis thaliana* contains ten cellulose synthases that exist in

multimeric complexes, Marchantia has only two [11]. Furthermore, since the dominant generation of *Marchantia* is the haploid gametophyte, heterozygosity can be circumvented to study mutant and transgenic phenotypes directly. Different genetic combinations can be obtained by crossing, where eggs produced by the archegonia are fertilized by motile sperm originating from antheridia [12]. Alternatively, clonal lines can be easily produced asexually by small disc-shaped propagules known as gemmae, which are found on the dorsal side of the haploid thallus [13]. Further advantages of *Marchantia* include the option to cryopreserve of gemmae [14] and spermatozoa [15], ~3 months sexual generation time, 2–3 weeks for asexual reproduction, and a small genome size (225.8 Mb, nine chromosomes)[8,9]. These features make *Marchantia* an attractive model to dissect the function of genes and biological pathways. Indeed, with >500 publications available on NCBI at the time of writing and with >50 publications in 2021 alone, Marchantia is a model organism on the rise.

Extreme environmental conditions can cause a multitude of biotic and abiotic stresses that can devastate crops and induce the collapse of entire ecosystems [16,17]. Plants have evolved sophisticated mechanisms to perceive and respond to the different stresses, which induces an acclimation process that allows the plant to survive the stress [18], but often at the cost of reduced growth [16,17]. Many studies have analysed the effect of stress on plant growth by identifying differentially expressed genes between stress-treated and untreated plants (e.g., [19–21]), or by identifying single nucleotide polymorphisms associated with stress resistance in *Arabidopsis* and maize [22–24]. The studies performed on non-agricultural species such as *Arabidopsis* can shed light on fundamental processes of stress acclimation, but it is currently unclear whether the acclimation mechanisms between different species or even accessions are conserved and transferable from model to crop.

In addition, plants are often exposed to a combination of stresses in the natural environment which can require opposing strategies for survival. For example, drought causes plants to close their stomata to minimize water loss [25,26], while heat causes the stomata to open to cool down their leaves via transpiration [21,27,28]. The stress signalling is mediated by a diverse ensemble of stress-specific sensors/receptors, networks of protein kinases/phosphatases, calcium channels/pumps, and transcription factors that can be localized to different organelles [29,30]. Stress signalling is further communicated and attenuated by hormones, other signalling molecules (e.g., reactive oxygen species) and protein modifications (s-nitrosylation, ubiquitination, myristoylation) [31,32]. Due to this complexity, it is challenging to elucidate the molecular basis of stress acclimation to one stress. The problem becomes even more intangible when multiple simultaneous stresses are considered. Apart from antagonistic reactions to combined stresses, it has been observed that a plant's response to different stresses is not equal and is dominated by the stress that is perceived to be more severe by the plant. This phenomenon has been observed when abiotic and biotic stresses intersect (refer to the

chapter by [33] for examples) and in combinations of heat and drought stress [34]. Therefore, a simpler model plant system, such as *Marchantia,* would be valuable in studying stress acclimation.

In this study, we set to take advantage of *Marchantia*'s less complex signalling architecture to understand better how plants respond to environmental cues such as stress and time of the day to modulate the expression of genes and biological pathways. To this end, we constructed an abiotic stress gene expression atlas of *Marchantia* comprising seven abiotic stresses (darkness, high light, cold, heat, nitrogen deficiency, salt, mannitol) and their pairwise combinations (e.g., cold + salt). We first compared the stress acclimation responses to other members of the plant kingdom to determine whether these responses are conserved across species. We then identified stress-responsive biological pathways and investigated whether new gene families evolved specifically to confer adaptations to abiotic stress. Next, we compared the transcriptomic responses of the different stress combinations (cross-stress) to identify the differentially regulated pathways and how the different stresses enhance or suppress one another. We then constructed a diurnal gene expression atlas and investigated whether diurnal gene expression is conserved between *Marchantia* and Archaeplastida. Finally, to make the data readily accessible to non-bioinformaticians, we provide i) an eFP browser for Marchantia that allows the visualization of gene expression in organs and stress conditions and ii) an updated CoNekT platform, allowing sophisticated comparative gene expression and co-expression analyses.

Material and Methods

Maintenance of *Marchantia polymorpha*

Male *Marchantia polymorpha*, accession Takaragaike-1 (kindly provided by Prof. Chen Zhong from National Institute of Education, Singapore) was propagated on half-strength Gamborg B-5 Basal agar (1% sucrose, pH 5.5, 1.4% agar)[35] in deep well plates (SPL Biosciences, SPL 310101) at 24°C under continuous LED light at 60 µEm^{-2}s^{-1}.

Experimental setup for stress experiment

In order to determine the ideal stress condition to be used for cross-stress experiments, a series of severity was used for heat, cold, salt, osmotic, light, dark, and nitrogen deficiency stresses. For each stress variation, three agar plates containing nine gemmae each were plated, where two plates were used as material for RNA sequencing and one plate was kept for observation until 21 days after plating (DAP).

For salt, osmotic, and nitrogen deficiency stress, the gemmae were plated on half-strength Gamborg B-5 Basal agar (pH 5.5, 1.4% agar, 12.4 mM KNO_3) supplemented with 20 - 200 mM NaCl (20 mM steps), 50-400 mM mannitol (50 mM steps) and KNO_3 concentration ranging from 90 to 0% (22.3 - 0 mM, 2.5 mM steps), respectively. The potassium concentration in nitrogen deficiency agar was maintained

using equimolar concentrations of KCl. Gemmae subjected to heat, cold, light, and dark stresses were plated on normal half-strength Gamborg B5 agar (pH 5.5, 1.4% agar, 12.4 mM KNO_3, 0.5 mM $(NH_4)_2SO_4$).

Plates were grown at 24°C under continuous LED light at 60 µEm^{-2}s^{-1} from days 0 to 13. For dark stress, plates were moved to the plant growth chamber (HiPoint M-313) on days 8, 9, 10, 11, 12, 13, and 14 for growth in darkness at 24°C for 7, 6, 5, 4, 3, 2 and 1 day(s), respectively. On day 14, all plates were transferred to the plant growth chamber (HiPoint M-313) for 24 hours. Control and plates for salt, osmotic, and nitrogen deficiency stresses were maintained at 24°C under continuous LED light at 60 µEm^{-2}s^{-1} in the plant growth chamber. Normal growth conditions were used for heat, cold, dark and light stresses with the following modifications. For heat and cold stress, 27°C to 36°C (3°C steps) and 3°C to 12°C (3°C steps) were used, respectively. For light stress, plants were subjected to 20% to 100% chamber capacity of light output (115 to 535 µEm^{-2}s^{-1}), at steps of 20%, approximating to 100 µEm^{-2}s^{-1} per step.

On day 15, whole plants were harvested by pooling three plants into 2 mL Eppendorf tubes, flash-frozen in liquid nitrogen, and stored at -80°C. This was done in triplicates, resulting in 3 (replicates) * 3 (pooled plants) used for each RNA-seq sample. Pictures of the front and back of the observation plates were taken, and the plates were returned to normal growth conditions. Pictures of the front and back of the observation plates were taken on day 21 without cover.

Cross stress experiments were carried out similarly with the following parameters for the various stress combinations - heat (33°C), cold (3°C), salt (40 mM NaCl), osmotic (100 mM mannitol), light (435 µMm^{-2}s^{-1}), darkness (3 days) and nitrogen deficiency (0 mM KNO_3).

Diurnal gene expression

For diurnal gene expression, three agar plates containing nine gemmae each were plated for each sampling time point on half-strength Gamborg B-5 Basal agar (0% sucrose, pH 5.5, 1.4% agar). The plants were grown at 24°C under continuous LED light at 60 µEm^{-2}s^{-1}. On day 8, plates were transferred to the plant growth chamber and grown at 24°C under LED light at 60 µEm^{-2}s^{-1} with a 12 h/12 h day/night cycle. On day 15, plants were harvested at Zeitgeber time (ZT) 2, 6, 10, 14, 18, and 22. The plants were stored in 2 mL Eppendorf tubes (3 plants per tube) and flash-frozen in liquid nitrogen before storage at -80°C.

Size measurement of *Marchantia polymorpha*

The photos of the observation plates were scaled and measured on Adobe Illustrator. Length and breadth of the thallus were taken with relation to the central axis (Figure 1B shows the plants with the central axis in a horizontal position). The thallus's approximate

area was derived by the product of the length and breadth of the thallus. Abnormally small plants (outliers) were excluded from further analysis.

RNA extraction and sequencing

Plants from each Eppendorf tube were ground into fine powder in liquid nitrogen using mortar and pestle. Total RNA was extracted using the Spectrum™ Plant Total RNA Kit (Sigma, STRN-250) using Protocol A (750 µL Binding Solution) according to manufacturer's instruction with on-column DNase digestion using 60 µL of DNase mixture (15 µL RQ1 RNase-Free DNase (Promega, M6101), 6 µL RQ1 DNase 10X Reaction Buffer and 19 µL nuclease-free water) per column.

Preliminary quality control of the extracted RNA (triplicates for each condition) was done using Nanodrop before further quality control checks by Novogene (Singapore) for sample quantitation, integrity, and purity using Nanodrop agarose gel electrophoresis and Agilent 2100 Bioanalyzer. Library construction from total RNA including eukaryotic mRNA enrichment by oligo(dT) beads, library size selection, and PCR enrichment was performed by Novogene using NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina®. The libraries were then sequenced with Illumina Novaseq-6000, paired-end sequencing at 150 base pairs and at sequencing depths of ~20 million reads per sample.

Expression quantification

RNA sequencing data were mapped against the *Marchantia polymorpha* CDS (v5.1 revision 1, MarpolBase), quantified, and TPM-normalised (transcript per million) using Kallisto v 0.46.1. Publicly available RNA sequencing data for *Oryza sativa* were downloaded from ENA as fastq files and mapped against the CDS of *Oryza sativa* v7.0 obtained from Phytozome v13 using Kallisto.

Identification of differentially expressed genes

Non-normalized counts from Kallisto were used for the analyses of differentially expressed genes (DEGs) using R package DESeq2 [36] where various stress conditions were compared against their respective controls.

For our *M. polymorpha* dataset, only genes that were found to be differentially expressed against controls from two different batches were considered for further analysis. For downstream analysis, only genes with an adjusted p-value < 0.05 and a −1> $\log_2$-fold >1 were considered as differentially expressed. DEGs of *Arabidopsis thaliana*, *Chlamydomonas reinhardtii,* and *Cyanophora paradoxa* were obtained from the study of [20].

Identification of transcription factors, kinases, and differentially expressed pathways

The biological function and pathway membership of genes of *Cyanophora paradoxa, Chlamydomonas reinhardtii, Marchantia polymorpha, Oryza sativa,* and *Arabidopsis thaliana* were annotated using Mercator 4 v2.0[37]. The annotation of transcription factors belonging to *C. reinhardtii*, *M. polymorpha*, *O. sativa*, and *A. thaliana* was retrieved from PlantTFDB v5.0[38], while the annotation of kinases was obtained from iTAK v18.12[39]. Transcription factors and kinases of *C. paradoxa* were inferred from orthologs of the four species identified by OrthoFinder (data from [40]).

Significantly differentially expressed pathways were determined through hypergeometric testing, where the observed number of DEGs in a pathway was compared to the permuted number of DEGs. The p-values were adjusted for multiple testing using Benjamini Hochberg correction (p-value < 0.05)[41].

Comparison of species and stresses with differentially expressed biological pathways and orthogroups

To estimate the similarities of species/stresses by the differentially expressed biological pathways, we determined for each species/stress the identity biological pathways showing significant enrichment upregulated (U), downregulated (D), both (UD) or not changed (N). Then, each experiment was represented as a vector of U, D, UD, or N values. The Jaccard distance between each experiment was calculated and used as the linkage for the cluster map. For a fair comparison of the experiments and species, biological pathways that were N (not changed) for both species/stresses were excluded. The modified Jaccard's distance between species/stress experiments X and Y is:

$$JD_{(X,Y)} = 1 - \frac{Gene\ family\ statuses\ in\ X\ \cap\ Gene\ family\ statuses\ in\ Y}{Gene\ family\ statuses\ in\ X\ \cup\ Gene\ family\ statuses\ in\ Y}$$

For the analysis where we used orthogroups as a basis for comparison, we used the same procedure, but with the following changes: for each species, an orthogroup was labeled as "None" if it did not contain any genes, "no change" if it did not contain DEGs (i.e., the expression was not affected by stress), "ambivalent" if it contained up and downregulated genes, and "upregulated" and "downregulated" for orthogroups that only contain up or downregulated genes, respectively. Then, each experiment was represented as a vector, where the items represented orthogroups with 'None', 'no change', 'ambiguous', 'upregulated' or 'downregulated' items. For a fair comparison of the experiments and species, gene families that are not present in either experiment (due to certain orthogroups not being present in all species), and the absence of DEGs from gene families in both experiments, were excluded.

The JD values range from 0 (the two species/stresses contain identical differentially expressed orthogroups or biological pathways) to 1 (the two species/stresses are entirely different). The JD values were used as linkages for clustering of the species/stresses.

Stress responsiveness of biological pathways and phylostrata

We compared the stress responsiveness of the orthogroups across various species and stresses by assigning gene families into five categories (as done above) for each stress: no genes found in family ('None'), no change in gene expression (NC, i.e., no DEGs), only upregulated DEGs are present (UP), only downregulated DEGs are present (DOWN), and both up- and down-regulated DEGs are present (AMB). We looked at the distribution of biological processes of differentially expressed genes that are found in the orthogroups that were responsive to stress. The counts were logged and row normalised by the total number of genes present in a Mapman bin.

A phylostratigraphic analysis was performed on the *Marchantia* stress perturbations and the corresponding stresses in the four other species. Through the statuses of the orthogroups assigned earlier, we considered an orthogroup to be stress-responsive if there were at least two of the same status (UP/DOWN/AMB) across species that are not None or NC. Alternatively, if there was a mixture of statuses, the orthogroup is assigned MIXED for that stress. The age of the orthogroup was determined by the presence of the oldest species present in our analysis. Similarly, species-specific orthogroups were assigned statuses based on their initial assignments. All other scenarios were considered as non-responsive and assigned the status of 'NR'.

Comparison of single and cross stresses with UpSet plot

Using the DEGs obtained from differential gene expression analysis, we looked at the top 50 intersections of all perturbations of *Marchantia* in terms of up and downregulated genes through an upset plot, provided by https://upsetplot.readthedocs.io/en/stable/.

Determination of rhythmic genes and biological pathways

Rhythmic genes were identified using JTK_Cycle v3[42] with parameters set at six timepoints with three replicates, temporal spacing of 4 hours between samples, and target period of 6 hours. Genes are considered expressed if the TPM value is more than 1 in at least one replicate. Genes are only considered rhythmic if it is supported by an adjusted p-value of < 0.05. The LAG value was corrected by two hours (i.e., LAG 0 to LAG 2) as our first timepoint is ZT2, meaning two hours after initial exposure to light.

To observe the rhythmicity of genes according to their biological function, we calculated the percentage of rhythmic genes that peak at the respective LAGs identified by JTK_Cycle for each Mapman bin. The linkage between the Mapman bins was determined by scipy hierarchical clustering.

Comparison of diurnal gene expression in *Marchantia* to other species

We compared the diurnal expression of 1-to-1 orthologs as done in [43]. JTK data from this publication included species *Cyanophora paradoxa, Chlamydomonas reinhardtii, Porphyridium purpureum, Klebsormidium nitens, Physcomitrella patens, Selaginella moellendorffii, Picea abies, Oryza sativa,* and *Arabidopsis thaliana.* In addition, we also looked at the number of orthologs that have the same phase in *Marchantia* and another species after applying a phase shift in the other species.

Data availability

The RNA-seq data capturing the expression of controls, single and double stresses are available from https://www.ebi.ac.uk/ena as E-MTAB-11141.

Script availability

Python and bash scripts used to generate the figures in the paper are available from: https://github.com/tqiaowen/marchantia-stress

Results

Subjecting *Marchantia polymorpha* to a combination of abiotic stresses

To capture gene expression changes caused by single stress or a combination of stresses, we first set to establish the type and magnitude of the different stresses. We defined two types of stresses: i) the environmental stresses comprised heat, cold, excess light, and prolonged darkness, while ii) media stresses comprised nitrogen deficiency, excess salt (representing ionic and osmotic stress), and excess mannitol (representing osmotic stress). Next, the magnitude of the stresses was modulated to identify sub-lethal stress conditions or growth decrease by ~50% (inferred from approximate thallus area). To this end, gemmae grown on sealed, sterile agar plates at 24 hours light were subjected to varying degrees of stress, and their phenotypes were observed on days 15 and 21. For media stresses, the gemmae were subjected to the stress from day 0, while for the environmental stresses, the stress was applied at day 14 for 24 hours (Figure 1A). Prolonged darkness was an exception to this design, as plates were subjected to darkness on days 8, 9, 10, 11, 12, 13, and 14 to expose the plants to 7, 6, 5, 4, 3, 2, and 1 day(s) of darkness at day 15, respectively (Figure 1A).

Single stresses showed varying degrees of effect on plant growth at day 15 (day of harvest, Figure S1 shows growth measurements, Supplemental Data 1 shows agar plates) and day 21 (6 days post-stress for environmental stresses, Figure 1B). One day

of cold stress did not affect the growth at the temperature range tested (3-12°C), and we selected 3°C for further analysis. The heat stress experiment showed that the plants abruptly died when the heat treatment temperature was increased from 33°C (no phenotype) to 36°C (death), and we selected 33°C for further study. For light stress, we selected 435 $\mu Em^{-2}s^{-1}$) as we observed necrosis at the next higher light intensity (535 $\mu Em^{-2}s^{-1}$, Supplemental Data 1). For osmotic (100 mM selected) and salt stress (40mM selected), we observed an expected negative growth gradient when the concentration of the two compounds was increased. For nitrogen deficiency, at 0 mM KNO_3, we observed a decrease in growth and an accumulation of a red pigment, which likely represents auronidin, a flavonoid shown to accumulate during phosphate deficiency [44]. Finally, for darkness, we observed that the growth of plants decreased proportionately with the duration of stress, and we selected plants grown in three days of darkness, as they showed a growth decrease of 50% (Figure S1). On day 21 (i.e., six days of normal growth condition), all dark-grown plants showed increased size, indicating that the seven days of darkness are not lethal.

Next, we set out to determine how *Marchantia* responds to a combination of two stresses. To this end, we tested all 19 possible pairs of stresses (cold+heat and dark+light combinations are mutually exclusive and hence, excluded) using the same experimental outline as for single stresses (Figure 1A). We did not observe any unexpected additive effects of any two stresses (Figure 1D), as typically, a combination of stresses resulted in an expected combination of phenotypes (e.g., nitrogen: small, pigmented, mannitol: small, nitrogen+mannitol: even smaller and pigmented) except for salt-nitrogen (SN), which is significantly larger than its salt counterpart but not different from its nitrogen counterpart (Figure 1E). While plants subjected to a combination of sub-lethal heat (33°C) and high light (435 $\mu Em^{-2}s^{-1}$) died (Figure 1D), this was most likely due to a greenhouse effect caused by high irradiation and sealed plates, as the temperature of agar climbed to 38°C (i.e., lethal temperature, Figure 1C). Interestingly, for high light and cold, we observed yellowing of the thalli, indicating that lowering the temperature makes the plants more sensitive to high light (Figure 1C).

Similar to the single stress experiments, the RNA of plants subjected to the combination of two stresses were isolated. The resulting panel of 7 single stresses, 18 combinations of two stresses and 2 untreated controls (controls from single stress and combined stress batches) were sent for RNA sequencing (Table S1 contains sample labels). The resulting data was used to estimate gene expression (Table S2, S3 contains Transcript Per Million (TPM) values and raw counts, respectively). Overall, we observed a good agreement between the sample replicates, as samples showed expected clustering (Figure S2), and the correlation between expression profiles of replicates was >0.97 (Table S4).

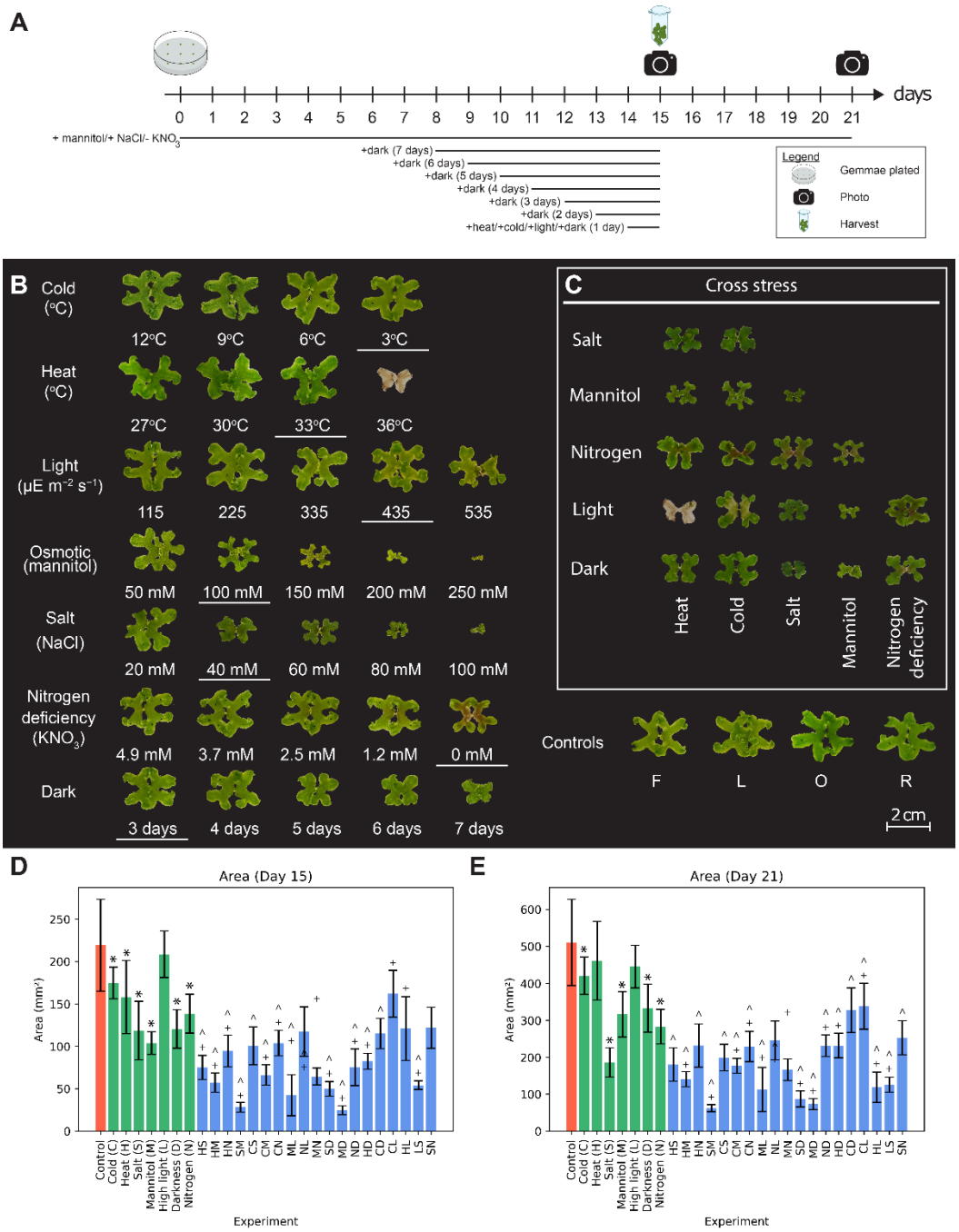


Figure 1. Influence of different abiotic stresses on the growth of *Marchantia polymorpha*. A) Overview of the experimental setup for stress experiments. Gemmae were plated on ½ Gamborg B-5 agar with necessary modifications relevant to the stress tested. Black lines below the timeline indicate the duration where plants were exposed to stress. All plates were moved to the plant growth chamber on day 14 at the growth conditions appropriate to the respective stresses. Plants were sampled on day 15, where three plants were pooled as a replicate sample for RNA-seq. Observation plates were returned to normal growth conditions on day 15, and photographs of the plates were taken on days 15 and 21. B) Phenotype of plants on day 21 for heat, cold, osmotic, salt, light, dark and nitrogen deficiency stresses. Conditions with underline

represent the stress intensities used for combined stress. Controls F and L are representatives from the batches at the start and end of independent stresses in panel B. C) Phenotypes of plants on day 21 for combined stresses. Controls O and R are representatives from the start and end of combined stresses in panel C. D and E) Thallus size of gemmalings on Day 15 and 21 of growth (1/7 day after physical stress, 15/21 days after chemical stress, error bars are represented by standard deviation and Student's two-tailed t-test, p-value < 0.05. Asterisk (*) represents significant difference to control, caret (^) represents significant difference to first single stress control and plus sign (+) indicates significant difference to second single stress control.

No evidence for conservation of abiotic stress responses in Archaeplastida

Model organisms, such as *Arabidopsis thaliana*, are studied with the aim to apply the gained knowledge to other plants, such as food crops. The usage of model organisms assumes that there is a degree of conservation of gene sequences, biological pathways, and how the different organisms respond to stress. To study the conservation of stress responses to abiotic stress across multiple members of Archaeplastida, we compared our stress expression data of *Marchantia polymorpha* (Mpo: heat, cold, high light, darkness, low nitrogen, salt, mannitol) with publicly available expression data for basally diverging alga *Cyanophora paradoxa* (Cpa: heat, cold, high light, darkness, low nitrogen, low sulphur, low phosphate, salt, mannitol), rice *Oryza sativa* (Osa: cold, salt, drought) and *Arabidopsis thaliana* (Ath: heat, cold, high light, darkness, low nitrogen, low sulphur, low phosphate, low iron, salt) (Table 1, Table S5)[45,46].

To assess how abiotic stress affects global gene expression, we first identified the percentage of differentially expressed genes (DEGs). Since our experiment was done in batches, we used two controls taken at the beginning and the middle of the experiment and set the requirement for a gene to show conserved differential expression in both controls to be deemed a DEG (Table S6 and S7). Overall, we observed a good agreement between the two controls, as both volcano plots (Figure S3) and set comparisons (Figure S4) indicated a similar set of DEGs. Standard differential gene analysis was also done for *O. sativa,* and a summary of the DESeq2 output and list of DEGs can be found in Tables S8 and 9, respectively. We observed that typically, only a small percentage of genes are differentially expressed, and the percentage of DEGs varies across species and stress, with dark stress having the most DEGs across all species (Figure 2A). The transcription factors (TFs) also comprised a small number of DEGs (Figure 2B), with the highest percentage in *A. thaliana* and *O. sativa* in general (list of TFs Table S10). The percentage of protein kinase DEGs follow the same trend as TFs, with the highest percentage in Arabidopsis and rice (Figure 2C) (list of kinases Table S11). In most experiments, there is an approximately even amount of up and downregulated genes except in *Arabidopsis* salt stress, where there is a disproportionate amount of downregulated genes and a larger proportion of upregulated genes in *Cyanophora* light, *Marchantia* mannitol and nitrogen, *Oryza* salt-1, and *Arabidopsis* low phosphate and low iron (Figure 2D).

Next, we wanted to understand which biological pathways are differentially expressed during stress and how these responses are conserved in Archaeplastida. To this aim, we first identified pathways that were significantly upregulated (more genes from the pathway are DEGs than expected by chance, adjusted p-value<0.05), downregulated, or both (a pathway contains more up- and down-regulated genes than expected by chance). This analysis was performed for each species and stress combination in our dataset, giving us stress/species (x-axis) and biological process enrichment (y-axis) matrix. The matrix was then subjected to hierarchical clustering, which grouped similar species/stress responses (by clustering columns) and similarly responding biological pathways (by clustering rows) (Figure 2E). Overall, we observed no apparent conservation between stresses in different organisms as the different stresses did not cluster well. At the same time, we did not observe any biologically meaningful clustering of biological pathways. For example, unrelated protein translocation pathways, chromatin organization and vesicle trafficking showed the highest similarity across all species/stresses (Figure 2E). To confirm these findings, we also investigated whether there are similarities in responses of gene families (Table S12). In this analysis, we assumed that if two species/stresses are similar, they should show a similar pattern of up- or down-regulated gene families. However, similar to the results for biological pathways, we observed no clear grouping of species/stresses (Figure S5), suggesting that gene expression responses are not conserved across species.

Table 1. Summary of stresses analysed across the five species in this study.

	Cpa	Cre	Mpo	Osa	Ath
Heat	37°C, 12 h (in darkness)	-	33°C, 24 h	-	44°C, 4h
Cold	4°C, 12 h (in darkness)	-	3°C, 24 h	10°C, 10h	10°C, 24 h
Osmotic	100 mM mannitol, 96 h	-	100 mM mannitol, 15 days	10% available water content (AWC), 3 days	-
Salt	75 mM NaCl, 96 h	-	40 mM NaCl, 15 days	300 mM NaCl, 24 h	150 mM NaCl, 13 days
High light	150 $\mu Em^{-2}s^{-1}$, 24 h	145 $\mu Em^{-2}s^{-1}$, 8 h	435 $\mu Em^{-2}s^{-1}$, 24 h	-	300 $\mu Em^{-2}s^{-1}$, 4 h
Prolonged darkness	0 $\mu Em^{-2}s^{-1}$, 3 days	-	0 $\mu Em^{-2}s^{-1}$, 3 days	-	0 $\mu Em^{-2}s^{-1}$, 2 days
Nitrogen deficiency	10% recommended KNO_3, 96 h	No nitrogen source, 2 days	0 mM KNO_3, 15 days	-	No nitrogen source, 2 days
Other nutrient deficiencies	Y	Y	-	-	Y

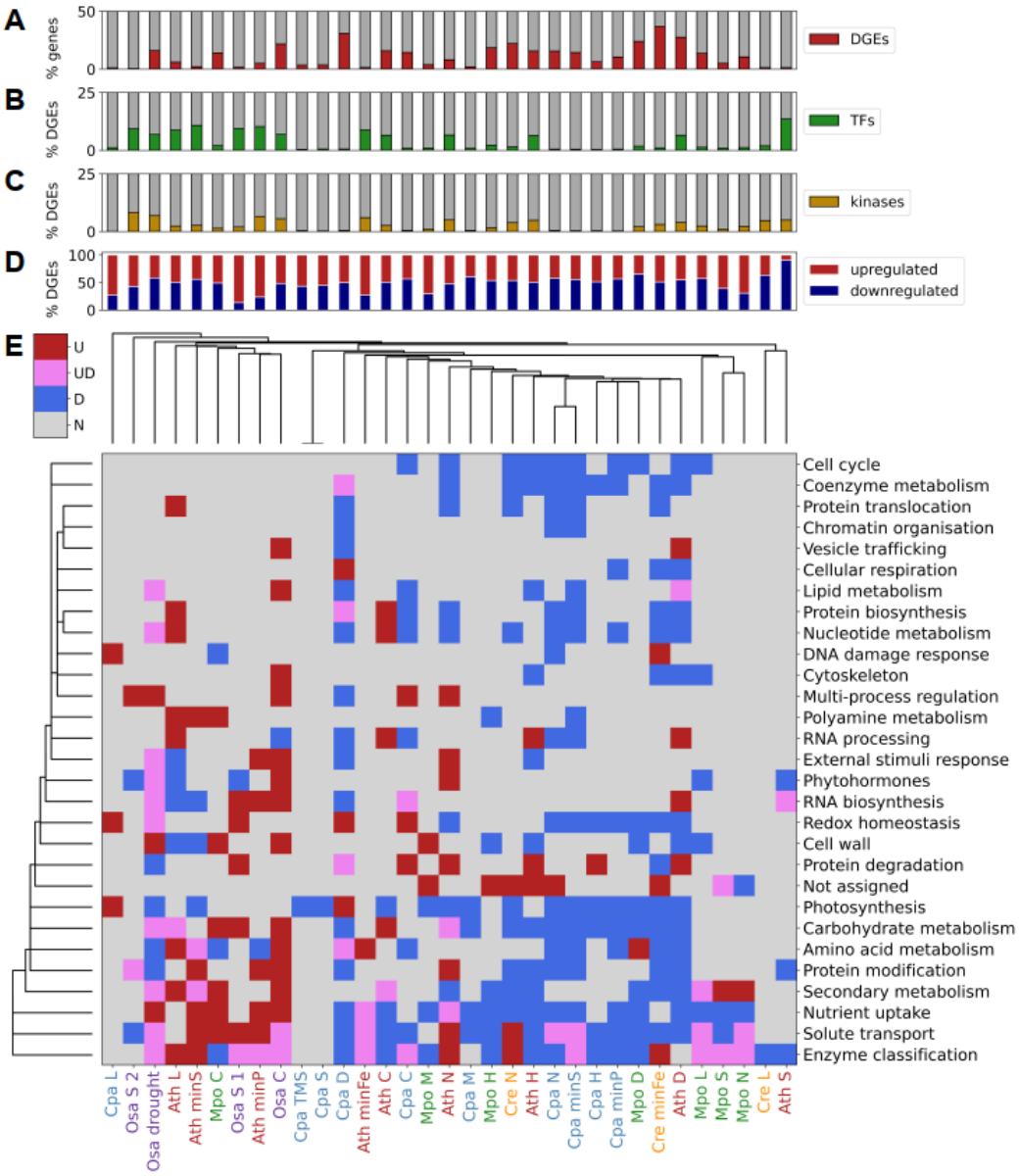


Figure 2. Comparative analysis of differential gene and pathway changes across species and stresses. A) Percentage of genes that are significantly differentially expressed. B) Percentage of differentially expressed transcription factors. C) Percentage of differentially expressed kinases. D) Percentage of up and downregulation of differentially expressed genes. E) Summary of biological pathways, represented by **M**apman bins, that are significantly enriched for differentially expressed genes. Colours in the heatmap, red, pink, blue**,** and gray correspond to significant upregulation (U), up and downregulation (UD), downregulation (D)**,** and no significant change (N), respectively. Experiment labels are coloured according to species, with blue, orange, green, purple**,** and red corresponding to *C. paradoxa*, *C. reindhardii*, *M. polymorpha*, *Oryza sativa**,*** and *A. thaliana**,*** respectively. The following abbreviations are used for the

description of stresses: cold (C), dark (D), heat (H), light (L), mannitol (M), nitrogen deficiency (N), salt (S), iron deficiency (minFe), sulfur deficiency (minS) and phosphorus deficiency (minP).

No evidence for the specialization of newly evolved gene families to stress acclimation

While gene expression response was not conserved in different species, the enrichment analysis in the previous figure revealed that specific biological pathways are more stress-responsive than others. For example, MapMan bin comprising 'Nutrient uptake' was more significantly differentially expressed across the different species/stresses than 'Vesicle transport' (Figure 2E).

To identify which biological pathways are highly responsive to the seven tested stresses in *M. polymorpha*, we first identified how many stresses a given gene family (represented by orthogroups) responds to. We observed that certain orthogroups are responsive to multiple stresses and have genes that are significantly upregulated (up; Table S13), downregulated (down), ambivalent (amb; some genes are up-, others are down-regulated), or not changed (nc) for one or more stresses. We then assigned the gene families to biological processes and plotted the stress responsiveness occurrences of the gene families (Figure 3A). Overall, we observed that gene families involved in DNA damage response, cell cycle, chromatin organization, protein biosynthesis/translocation are typically responding to one or two abiotic stresses, suggesting that these pathways are typically not affected by abiotic stress. Conversely, gene families involved in various uncharacterized enzymatic reactions ('Enzyme classification'), transcriptional gene regulation ('RNA biosynthesis') and protein degradation, cell wall, solute transport, and nutrient uptake were frequently differentially regulated during up to seven abiotic stresses, suggesting that these pathways are highly responsive to abiotic stress (Figure 3A).

Recently established gene families could be important for adaptation to the local environment. To test this hypothesis, we investigated whether there is a dependency between the age of a gene family and its stress responsiveness. First, by using phylostratigraphy (Table S14[47]), we assigned gene families to eight phylostrata. We sorted them from oldest to youngest (archaeplastida:Plantae, viridiplantae: green plants, embrophyta: land plants), and the species-specific phylostrata Mpo (*Marchantia*), Cpa (*Cyanophora*), Ath (*Arabidopsis*), Osa (*rice*), and Cre (*Chlamydomonas*). We then plotted the number of gene families, and the number of abiotic stresses they respond to and color-coded them by their phylostrata (Figure 3B). For all phylostrata, we observed a clear negative relationship between the number of gene families (y-axis) and the stress responsiveness (x-axis), indicating that most gene families typically respond to at most one or two stresses. Interestingly, the plots for the different phylostrata were largely parallel (e.g., Archaeplastida and Marchantia-specific lines are parallel and

overlapping), suggesting that there is no dependence between the age of gene families and their age. The only exceptions to this trend were rice- (Osa) and Chlamydomonas-specific (Cpa) gene families that showed a more negative slope, but this can be explained by these species only having data for two and three stresses, respectively.

We conclude that there is no dependency between the age of gene families and their responsiveness to stress. Thus, more recent gene families have likely not evolved to help plants acclimatize to abiotic stress.

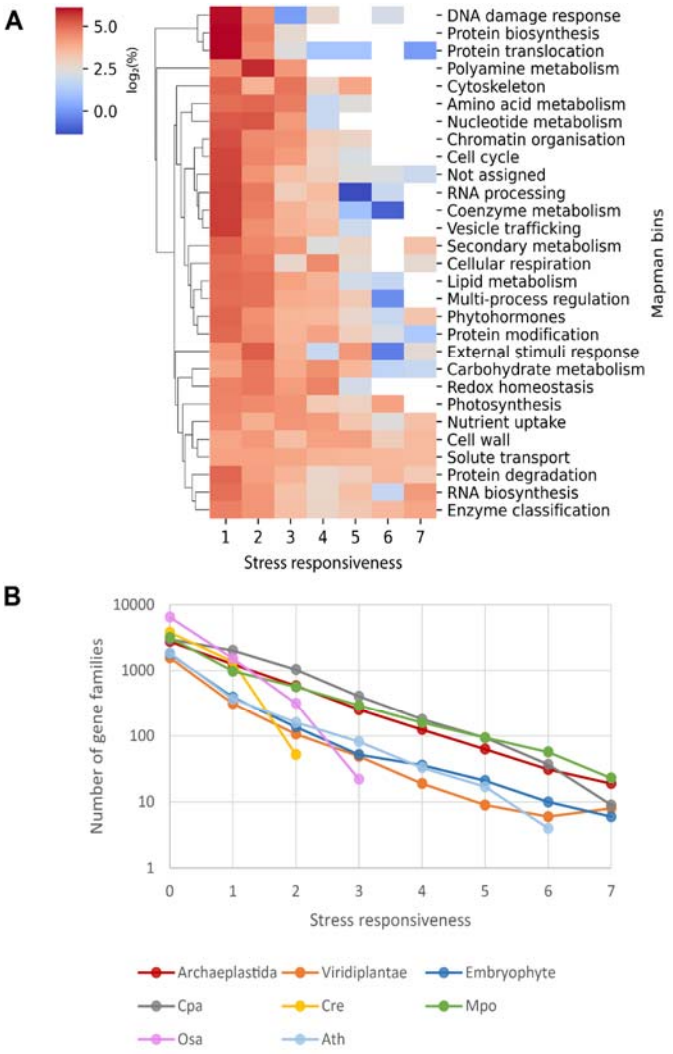


Figure 3. Analysis of stress responsiveness, biological function, and age of gene families. A) The proportion of functionally annotated orthogroups responding to the seven single (e.g., cold, heat) abiotic stresses in *Marchantia* and the equivalent stresses in the four other species. The data was row-normalized to show the proportion of the orthogroups responding to the number of abiotic stresses. B) Number of orthogroups found in phylostrata at various levels of stress responsiveness.

Combinatorial analysis reveals the hierarchy of stress responses in Marchantia

Plants experience many stresses when growing in nature, and these stresses need to be perceived and responded to. To better understand how Marchantia responds to single stress or a combination of two stresses, we first identified differentially expressed genes in the 18 combinations of two stresses and then compared the number of up- and down-regulated genes with the single stresses. We observed that the combination of stresses typically contains a similar or higher number of DEGs when compared to single stresses (Figure 4A). Next, to see if a severe growth phenotype results in a large number of differentially expressed genes, we plotted the plant size (x-axis) and the number of differentially expressed genes (y-axis, Figure 4B). We observed no significant correlation between these two variables for plants at days 15 and 21 (p-value > 0.05, Figure 4B). We concluded that there is no correlation between the severity of growth phenotype and the transcriptomic response to stress, as, e.g., the smallest plants (salt+mannitol, SM) also had the fewest number of differentially expressed genes.

The different single and combined stresses are likely to elicit similar and unique gene expression responses, resulting in the stresses having similar sets of DEGs. To elucidate these similarities, we used the UpSet plot, which shows the intersection of multiple sets for up-regulated (Figure 4C) and down-regulated (Figure 4D) genes. Interestingly, the largest set of up- and down-regulated genes was unique to heat+darkness combined stress (HD), suggesting that HD elicits the most unique and dramatic transcriptional response among the tested stresses. Other unique stress responses comprised upregulated cold+darkness (CD), cold+high light (CL), cold+nitrogen deficiency (CN) and heat+nitrogen deficiency (HN, Figure 4C); and downregulated cold+darkness (CD), cold+nitrogen deficiency (CN), cold (C) and cold+high light (CL). Darkness alone (D) and in combination with other stresses (e.g., ND, CD, MD, SD, HD) also contained a high number of up-regulated (connected dots in column 4,5, Figure 4C) and down-regulated (column 2,4, Figure 4D) genes, suggesting a conserved, core darkness response. Similarly, we observed core responses to heat (e.g., 6th column, Figure 4C) and cold (10th column, Figure 4C). Interestingly, we also observed DEGs across heat and darkness experiments (Figure 4C, D), suggesting that these two stresses elicit a similar response to a degree.

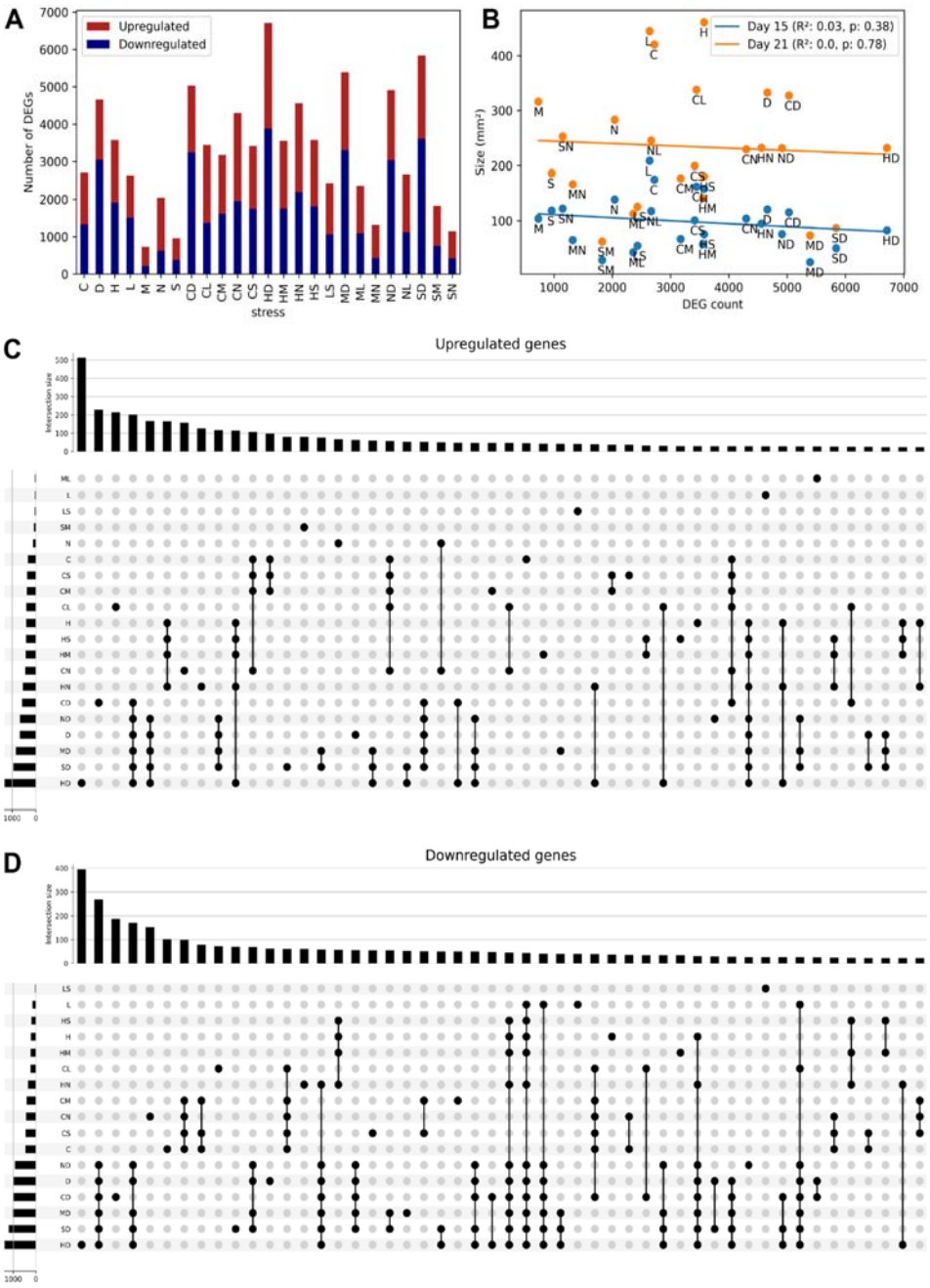


Figure 4. The number of differentially expressed genes in the tested stress combinations. A) The number of significantly (adjusted p-value < 0.05 against two controls) upregulated (red) and downregulated (blue) differentially expressed genes. B) The number of DEGs (x-axis) versus the size (y-axis) of Marchantia plants at day 15 (blue) and day 21 (orange). The R^2 and p-values are shown in the legend. Upset plot showing top 50 intersections of C) upregulated genes and D) downregulated genes. The bar charts on the top show the number of DEGs in the intersections, while the bar charts on the left show the number of DEGs in a given stress combination. The back dots indicate a given stress combination, while vertically connected dots indicate which

stress combination DEGs comprise a given intersection. For example, the 4th column in C) indicates that CD, ND, D, MD, SD**,** and HD stresses share ~200 upregulated DEGs.

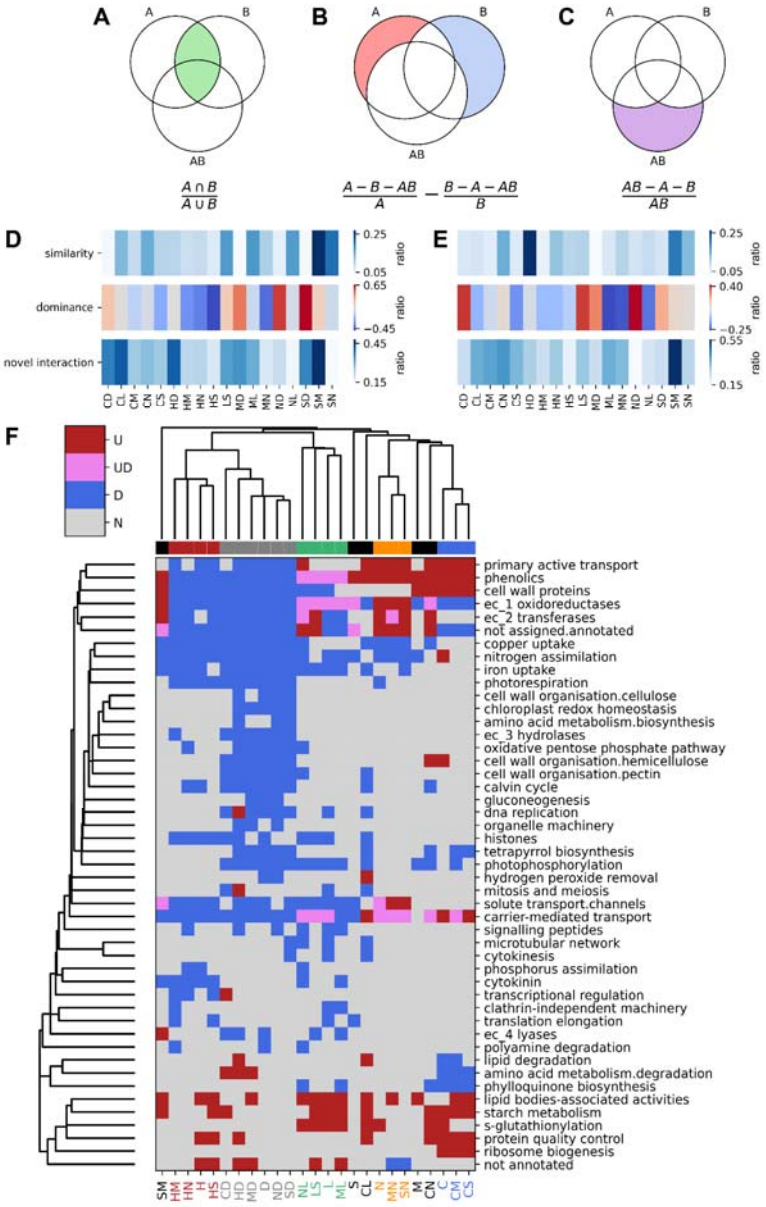
To better understand how Marchantia responds to a combination of two different stresses, we compared the combined response (AB) to the response to individual stresses (A and B), with three different metrics measuring the shared response, dominance of a stress and novel responses induced by combined stress. We first produced Venn diagrams for upregulated (Figure S6) and downregulated (Figure S7) gene sets. The first metric measures the similarity between A and B (Figure 5A, green area, Jaccard Index) and ranges from 0 (A, B do not have any DEGs in common) to 1 (A, B elicit identical DEGs). The second metric measures whether one stress suppresses the other (Figure 5B, the difference between red and blue area) and ranges from -1 (AB is similar to A but not to B, i.e., A suppresses B) to 1 (AB is similar to B but not A, i.e., B suppresses A). The third metric measures whether a combination of two stresses elicits a different, novel response when compared to the two individual stresses (Figure 5C, purple area specific to AB) and ranges from 0 (all DEGs are found in A and B, i.e., no novel response) to 1 (all DEGs in AB are unique, i.e., the combination of AB elicited a unique response).

Stresses showing the highest similarity in terms of DEG responses comprise salt and mannitol for up- and downregulated genes (Figure 5D-E, dark area for SM), salt and nitrogen deficiency for upregulated genes (Figure 5D), and heat and darkness for downregulated genes (dark area for HD, Figure 5E). The suppression analysis showed that darkness is a strong suppressor for many stresses (red boxes in CD, MD, ND, SD), except HD, indicating that heat stress and darkness are comparably dominant (Figure 5D-E). To support this observation, heat stress could also suppress other stresses (blue boxes in HS, HN, HM, Figure 5D-E). Finally, the novelty analysis revealed that the salt+mannitol combination elicited DEGs that were not found in the individual stresses (dark blue boxes for SM, Figure 5D-E), suggesting that the two stresses can activate new responses when combined.

To better understand the hierarchies of the stresses and how these stresses affect the activity of biological processes, we identified which processes contain significantly more upregulated or downregulated DEGs than expected by chance for each stress combination (Figure 5F). The clustering of the rows (biological processes) and columns (stresses) revealed that darkness-containing stresses form a clear group of similar responses (six stress combinations comprising D, HD, CD, MD, SD and ND), confirming the previous observation of the darkness suppressing other stresses. Interestingly, the darkness caused a dramatic decrease in gene expression of numerous pathways (Figure 5F, blue squares). The next largest group contained nearly all heat stress combinations (four stresses H, HD, HS, HN) and a similar but less dramatic decrease in the activity of many biological processes. Interestingly, despite the dramatic downregulation of biological processes in most dark and heat stresses, a

subset of stresses (H, HS, HD, MD) was significantly upregulated for uncharacterized genes (bin not annotated), suggesting that these responses employ poorly understood genes. The other groups comprised high light (four stresses L, NL, LS, ML), nitrogen deficiency (three stresses N, MN, and SN), and cold (three stresses C, CM, and CS). In contrast, salt- and mannitol-containing stresses did not form any groups, suggesting that despite dramatic phenotypic changes, these stresses are suppressed by other stresses (Figure 5F, M and S are not grouped with other stresses). Interestingly, salt+mannitol, cold+high light, and cold+low nitrogen were also not grouped, indicating that these combinations result in novel transcriptomic responses.

Taken together, we can rank the strength of dominance of abiotic stresses as darkness (six stress combinations), heat and light (four each), nitrogen deficiency and cold (three each), salt and mannitol (none).


A
B
C
D
E
F
similarity
dominance
novel interaction
ratio
U
UD
D
N
primary active transport
phenolics
cell wall proteins
ec_1 oxidoreductases
ec_2 transferases
not assigned.annotated
copper uptake
nitrogen assimilation
iron uptake
photorespiration
cell wall organisation.cellulose
chloroplast redox homeostasis
amino acid metabolism.biosynthesis
ec_3 hydrolases
oxidative pentose phosphate pathway
cell wall organisation.hemicellulose
cell wall organisation.pectin
calvin cycle
gluconeogenesis
dna replication
organelle machinery
histones
tetrapyrrol biosynthesis
photophosphorylation
hydrogen peroxide removal
mitosis and meiosis
solute transport.channels
carrier-mediated transport
signalling peptides
microtubular network
cytokinesis
phosphorus assimilation
cytokinin
transcriptional regulation
clathrin-independent machinery
translation elongation
ec_4 lyases
polyamine degradation
lipid degradation
amino acid metabolism.degradation
phylloquinone biosynthesis
lipid bodies-associated activities
starch metabolism
s-glutathionylation
protein quality control
ribosome biogenesis
not annotated

605

Figure 5. Comparison of differentially expressed genes and biological pathways in single and combined stresses. Illustration and equation of metrics used to measure A) similarity between two stresses, B) suppression of one stress when two stresses are combined, and C) novel genes that are differentially regulated when stresses are combined. Interaction between independent and cross stresses for D) upregulated genes and E) downregulated genes. In each stress combination, the first ('A') and second ('B') stress is the first and second letter of the combined ('AB') stress, respectively. For the suppression measure, a darker shade of red and blue indicates that more genes from the first ('A') and second ('B') stress are not represented in the combined stress, respectively. F) Heatmap of biological pathways (captured by second-level Mapman bins) that are significantly differentially expressed (hypergeometric test, adjusted p-value < 0.05). For brevity, we only show Mapman bins that are differentially expressed in at least three stress perturbations. Arbitrary clades of stress perturbations are highlighted by the different colours red, grey, green, orange, and blue squares at the tip of the dendrogram, and the leaf labels are also coloured. Abbreviations used to describe the categories of regulation are upregulation ('U', red), up and downregulation ('UD', purple), downregulation ('D', blue), and no change ('N', gray).

Comparative analysis reveals the conservation of diurnal gene expression between Marchantia and other Archaeplastida

Most members of the Archaeplastida kingdom subject more than one-third of their genes to diurnal gene expression, underlining the importance of diurnal regulation[43]. To study diurnal gene expression in Marchantia, the plants were grown in a 12 hour light / 12 dark cycle, and six timepoints were collected. The timepoints are 2, 6, 8 hours after light on and 2, 6, 8 hours after the light was switched off, corresponding to zeitgeber timepoints ZT2, 4, 6, 8, 12, 16, respectively.

The resulting 18 RNA-seq samples (six timepoints in triplicates) revealed the expected circular PCA plot (Figure S8A), and the gene expression from the individual samples showed an expected phasing pattern (Figure 6A). We identified 51.1% genes with significantly rhythmic gene expression (adjusted p-value < 0.05) with Jonckheere-Terpstra-Kendall (JTK_Cycle) algorithm[42], which revealed the time of the day where each rhythmic gene peaked (phase, Table S15). We also observed an expected circular pattern of the PCA plot of the genes (Figure S8B). The distribution of phases showed two peaks at ZT10 (i.e., 10 hours after lights on) and ZT20 (i.e., 4 hours before lights on, Figure 6B), which was reminiscent of phase distribution in Arabidopsis and rice[43]. Our previous study indicated that cell division genes are peaking just before the light / dark change in single cellular algae *Cyanophora paradoxa*, *Porphyridium purpureum*, *Chlamydomonas reinhardtii* and simple multicellular algae *Klebsormidium nitens*[43], showing that the diurnal rhythm regulates cell division in these algae. In contrast, Marchantia showed a more uniform expression of cell division markers (Figure 6C), indicating that, similarly to other land plants, Marchantia's cell division is not under the control of diurnal rhythm.

To elucidate how the different biological pathways are active during a typical day of Marchantia, we visualized the phase distribution of rhythmic genes and their pathways and performed a clustering analysis to reveal which pathways show similar diurnal gene expression (Figure 6D). Not surprisingly, we observed that, e.g., photosynthesis genes peak in expression during the day (Figure 6D, photosynthesis is labeled green during ZT 0-10) and shows a similar pattern with redox homeostasis, which is important for scavenging harmful reactive oxygen species generated by light capture [48]. Interestingly, and in line with the phase distribution plot (Figure 6A), we observed that most biological pathways tend to peak at two times of the day: during the day at ~ZT10 and the night at ~ZT20.

Our previous study showed that diurnal gene expression is conserved across >1 billion years of plant evolution[43]. To investigate whether Marchantia's diurnal gene expression is also conserved, we compared the distribution of phases of Marchantia genes and their orthologs in four algae (*Cyanophora paradoxa*, *Porphyridium purpureum*, *Chlamydomonas reinhardtii,* and *Klebsormidium nitens*) and five land plants (*Physcomitrella patens, Selaginella moellendorffii, Picea abies, Oryza sativa,* and *Arabidopsis thaliana*). The analysis plots the phases of Marchantia genes (y-axis, phases are arranged from 0 (top) to 23 (bottom)) versus the phases of the 1-to-1 orthologues from other species (x-axis, phases are arranged from 0 (left) to 23 (right)). Two species with similar diurnal gene expression would show an accumulation of points at similar x and y coordinates, indicating that the orthologs of the two species peak at the same time of the day. Conversely, the points would be randomly distributed in the plot for two species with different diurnal gene expressions. The analysis revealed that for most species, with the exception of Selaginella and Picea, we observe groups of genes at similar x and y coordinates, suggesting that the diurnal gene expression of Marchantia is conserved (Figure 6E). To further quantify the conservation of diurnal gene expression, we plotted the distribution of the differences between phase values of one-to-one orthologues of Marchantia and the nine Archaeplastida. For all species, except for Selaginella, Picea, and Arabidopsis, we observed a clear peak at around 0 (Figure 6F), confirming that the diurnal gene expression of Marchantia is conserved to most Archaeplastida. Taken together, the conservation between Marchantia and the other Archaeplastida makes Marchantia an attractive model to study diurnal gene expression.

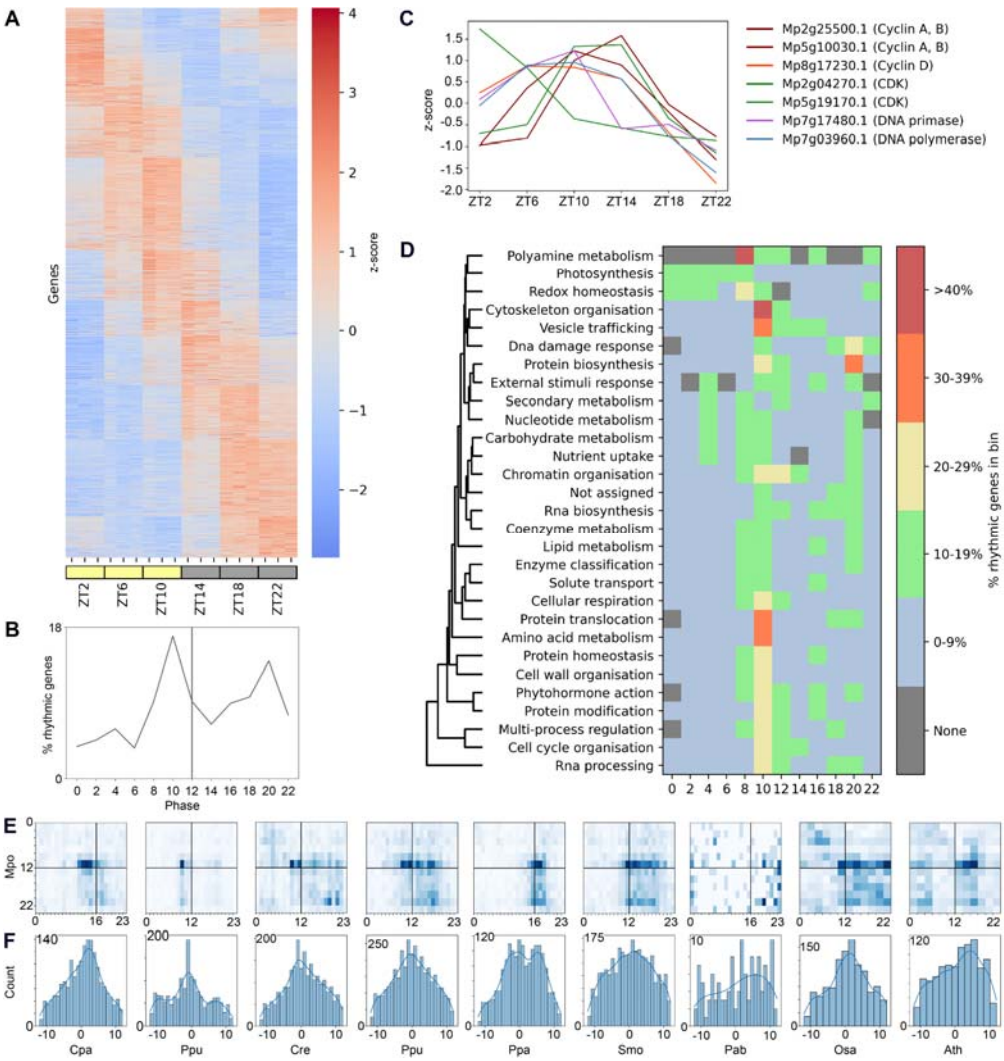


Figure 6. Comparison of diurnal rhythms of Marchantia and other Archaeplastida. A) Z-score-normalised expression of genes (y-axis) taken at different zeitgeber timepoints (x-axis). Yellow and grey boxes indicate the presence and absence of light at the time of sampling, respectively. The ticks indicate the biological replicates. The data was collected in triplicates. B) Distribution of rhythmic genes (as %, y-axis) across phases (x-axis). C) Z-score-normalised expression of significantly rhythmic (adjusted p-value < 0.05) cell cycle genes (cyclin A, B, D, cyclin-dependent kinase (CDK), DNA primase, and DNA polymerase). The y-axis and x-axis indicate the gene expression and average expression at given ZT, respectively. D) Diurnal phase peak distribution of biological pathways. Rows represent biological pathways, while the columns indicate the phases at which the genes in a pathway peak. The colors of the cells indicate the % of rhythmic genes in a given pathway at each phase. The sum of each row is 100%. E) The heatmap shows the phase distribution of phases of Marchantia genes (y-axis, phases are arranged from 0:top to 23:bottom) versus the phases of the 1-to-1 orthologs from other species (x-axis, phases are arranged from 0:left to 23:right). The intensity of

the cells in the heatmap **is** proportional to the number of genes that peak at a given phase for the two species. Black lines indicate the light to dark transition. Abbreviations used to represent species are as follows: *Cyanophora paradoxa* (Cpa)*, Chlamydomonas reinhardtii* (Cre)*, Porphyridium purpureum* (Ppu)*, Klebsormidium nitens* (Kni)*, Physcomitrella patens* (Ppa)*, Selaginella moellendorffii* (Smo)*, Picea abies* (Pab)*, Oryza sativa* (Osa) and *Arabidopsis thaliana* (Ath). F) Distribution of differences of phases of the 1-to-1 orthologues between Marchantia and the nine members of Archaeplastida. Value of 0 indicates that the 1-to-1 orthologs peak at the same time of the day**; the** value of -10 and 10 shows that the orthologs peak 10 hours earlier and 10 hours later in other species than in *Marchantia*, respectively.

eFP browser and CoNekT database for Marchantia

Bioinformatic data is only as useful as its accessibility. To make our data readily accessible, we have constructed an eFP browser instance for Marchantia available at https://bar.utoronto.ca/efp_marchantia/cgi-bin/efpWeb.cgi [49], and updated our CoNekT database [50], with our stress and diurnal data, available at https://conekt.sbs.ntu.edu.sg/. To exemplify how our data and these databases can be used, we provide an example with phenylpropanoids, which contribute to all aspects of plant responses towards biotic and abiotic stimuli. In flowering plants, phenylpropanoids were found to be highly responsive to light or mineral treatment, important for resistance towards pests[51], and invasion of new habitats and reproduction[52]. The biosynthesis of phenylpropanoids begins with phenylalanine ammonia lyases (PAL) and tyrosine ammonia lyases (TAL) that catalyze the non-oxidative deamination of phenylalanine to trans-cinnamate, which directs the output from the shikimate pathway phenylpropanoid metabolism[51].

To study phenylpropanoid metabolism in Marchantia, we started with entering 'PAL1' into CoNekT's search box (https://conekt.sbs.ntu.edu.sg/), which took us to the page of PAL1, a PAL gene *AT2G37040* from *Arabidopsis thaliana*. To identify PAL genes in Marchantia, we clicked on the link of the Land Plants orthogroup (CoNekT provides orthogroups for Archaeplastida, Land Plants, and Seed Plants), which revealed that all 11 land plants in the database contain PALs, while the algae *Cyanophora paradoxa* and *Chlamydomonas reinhardtii* do not. Marchantia contains a surprisingly large number (13) of PALs given the low redundancy of its genome, which is higher than Arabidopsis (4). CoNekT contains gene trees that also show the expression of genes in major organ types, and the analysis revealed that the many PAL genes in Marchantia are likely a result of a recent duplication within Marchantia (Figure S9).

To gain insight into the function of the 13 PAL genes from Marchantia, we set out to study the expression of the PALs during stress conditions. We copied the gene identifiers into Tools/Create heatmap (https://conekt.sbs.ntu.edu.sg/heatmap/), revealing that most PALs show a high expression during cold treatment combined with

nitrogen starvation (Figure 7A). We clicked on one of the highly responsive genes *Mp4g14110.1*, which took us to the page dedicated to the gene. The page contains various information, such as gene description, CDS/protein sequences, gene families, found protein domains, co-expression networks, and others. The detailed expression profile confirmed the high cold + nitrogen-specific expression of *Mp4g14110* by using CoNekT expression profile plot (Figure 7B) and eFP viewer (Figure 7C, red square at intersection of cold and nitrogen deficiency). To better understand the function of *Mp4g14110.1*, we clicked on the cluster link that directed us to the co-expression cluster link of *Mp4g14110.1*. The cluster page contains information about the genes found in the cluster, Gene Ontology enrichment analysis, found protein domains, and other functional information. Interestingly, the 'Similar Clusters' cluster page revealed similar co-expression clusters in other species (similarity is based on ortholog membership and defined by Jaccard Index), with another similar cluster in Marchantia. To study these duplicated clusters, we clicked on the 'Compare' button, which revealed that the two clusters contain several gene families involved in phenylpropanoid biosynthesis (yellow rounded rectangles: PALs, brown rectangles: chalcone synthases), ABC and DTX transporters implicated in metabolite transport across membranes (purple/gray/green and red rectangles[53]), auresidin synthases that can hydroxylate or cyclize chalcones (light blue rounded rectangles[54]) and glutathione transferases (red rounded square[55]). Interestingly, both clusters contained WRKY transcription factors (salmon rectangles), implicating these transcription factors in controlling the biosynthesis of the respective metabolites.

Taken together, our tools revealed evidence of duplicated modules, likely involved in the biosynthesis of related phenylpropanoids. The updated CoNekT platform contains many additional tools to predict gene function and find relevant genes [40,50,56]. For example, the tool found many of the described genes by identifying genes highly expressed during combined cold and nitrogen starvation (Table S16), by clicking on Tools/Search Specific Profiles, selecting Marchantia and ‘Cold-Nitrogen deficiency’.

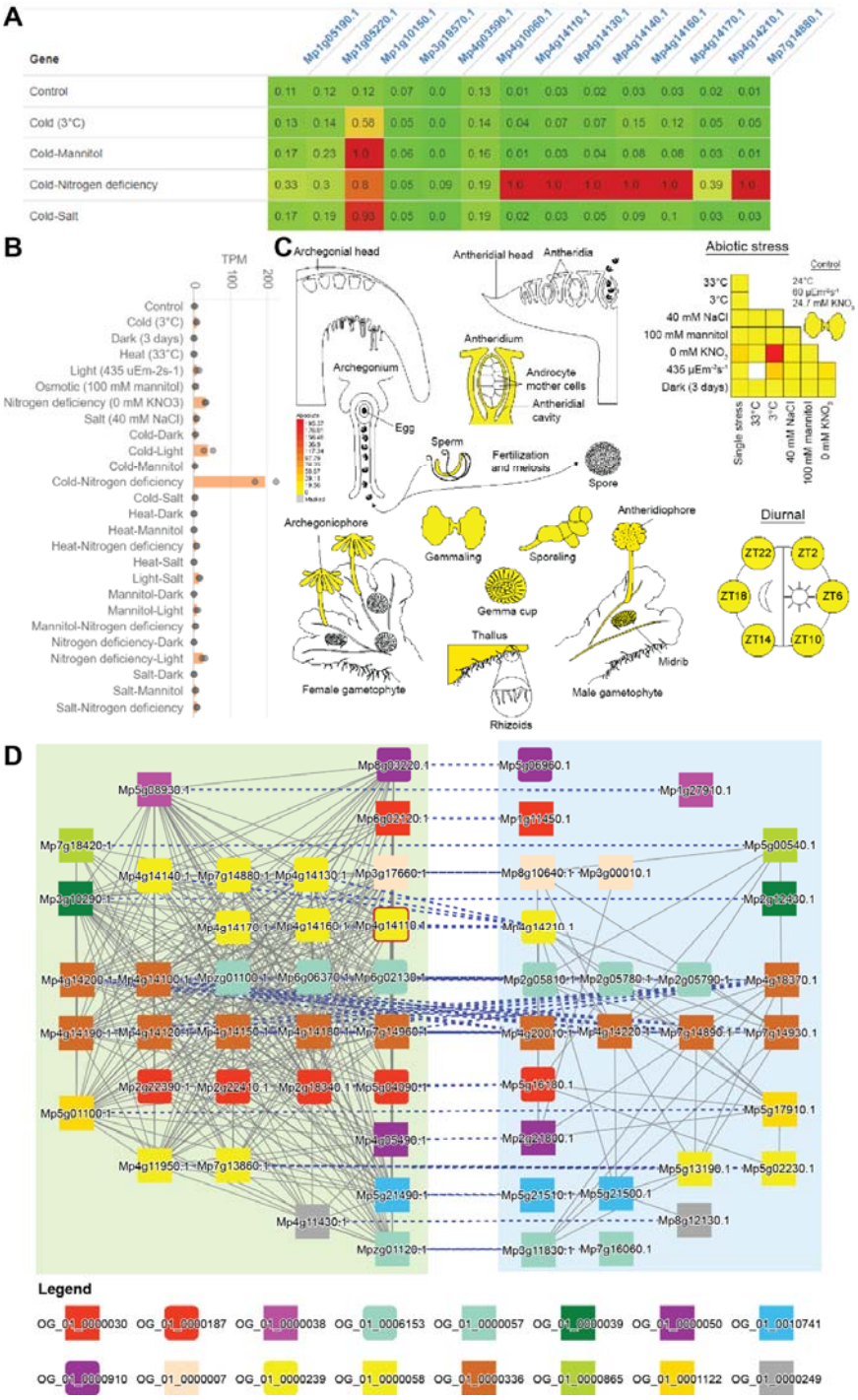


Figure 7. Implementation of the Marchantia gene expression data in CoNekT and eFP browser. A) Heatmap showing expression of PALs in representative stresses. B) Expression profile of Mp4g14110.1 (*MpPAL7*) under single and combined stress in CoNekt. C) Expression of *MpPAL7* in different organs and under different abiotic stress conditions in eFP. D) Comparison of Cluster 102 and 50 bounded by green and blue box, respectively. The border of node MpPAL7 is coloured red and blue dashed lines indicate homology between genes from the two clusters.

Discussion

To study how plants perceive changes in their environment and how conserved their responses are, we constructed a diurnal and stress expression atlas capturing gene expression changes to single and combined stresses for *Marchantia polymorpha*. With this data in hand, we set out to investigate how conserved the responses to the environmental changes are between model plants and which genes and pathways are responding to single or combined stress.

To determine the response of *M. polymorpha* towards various stresses, we tested a range of severity for heat, cold, salt, osmotic, light, dark, and nitrogen deficiency on *M. polymorpha* gemmae. As expected in most stresses, the size of the plants decreased proportionally to the severity of stresses (Figure 1B). Certain stresses, such as 3°C cold and 535 uEm^{-2}s^{-1} light, caused only minor growth phenotypes, suggesting that *Marchantia* can survive under even lower temperatures and higher light intensities. Conversely, stresses such as osmotic (100mM mannitol), salt (40mM NaCl), and carbon starvation (3 days of darkness) produce strong growth phenotypes resulting in reduced growth (osmotic, salt, darkness) and discoloration (yellow-green for mannitol > 150mM, darker plants for salt > 60mM salt) (Figure 1B). Death of plants occurred for heat stress at 36°C for 24 hours and at higher salt concentrations (>200mM). Interestingly, the plants grew at 0mM KNO_3, albeit slower and with reddish discoloration likely caused by auronidin, a flavonoid shown to accumulate during phosphate deficiency[44].

In most cases, the combination of stresses resulted in additive phenotypes (e.g., salt+mannitol stress results in smaller plants than the two stresses separately, Figure 1C-E). While heat (33°C) and high light intensity (435 uEm^{-2}s^{-1}) resulted in death, this was caused by temperature built up in the sealed plates, causing the temperature to rise to the lethal 38°C. Consequently, this stress combination should be performed using an open plate setup in the future. The only exception to this was salt+nitrogen deficiency (SN) (40 mM NaCl, 0 mM KNO_3), which was significantly larger than plants exposed only to salt stress and was observed to have extensive and dense rhizoids (Supplemental Data 1). Curiously, we did not observe any dramatic phenotypes when combining carbon/energy starvation (3 days of darkness) with any other stresses (Figure 1C). This is counterintuitive as, e.g., heat stress acclimation is a costly process requiring the biosynthesis of new transcripts and proteins to repair, replace and rebalance the affected cellular machinery[57]. A hypothesis for the unexpectedly mild phenotype for combined dark stress could be due to the lack of photosynthetic processes in the absence of light, which increases the plant's capacity to cope with increased ROS, a common response to most abiotic stresses[58]. This is also exemplified by the significant downregulation of oxidoreductases and chloroplast redox homeostasis in combined dark stress (Figure 5F). These phenotypes demonstrate that Marchantia is

able to survive various adverse growth conditions and can serve as an excellent model to study stress acclimation.

Worryingly, we observed no evidence for conservation of global stress acclimation responses between any of the five species of algae (*Cyanophora* and *Chlamydomonas)* and land plants (*Marchantia*, rice and *Arabidopsis)*, when comparing the responses using biological pathways and gene families (Figure 2E, Figure S5, respectively). While surprising, this is not entirely unexpected as massive changes such as genome and gene duplications have occurred over time, rendering a lack of homology in genes across species and differences in gene families and regulation. This suggests that each of the model plants uses a different strategy to acclimatize to stress, making it uncertain to what degree knowledge gained from model species such as *Arabidopsis* can be used to improve our crops. This lack of conservation of responses to abiotic stresses has been observed by us at the species level in *Cyanophora*[20], and at the intraspecies level by a salt stress study in six Lotus accessions, where only 1% of genes showed a conserved response[59], in seven *Arabidopsis* accessions, which showed a divergent response to treatment by salicylic acid[60] and by two strawberry cultivars, which displayed modest conservation of DEGs to the same pathogen[61]. However, since mRNA levels are only capturing one layer of gene regulation, the lack of conservation of responses to the abiotic stresses should be further investigated with other omic technologies measuring, e.g., protein levels, translation rate, transcription rate, and others.

Our analysis revealed that gene families comprising various enzymes (MapMan bin Enzyme classification), mRNA expression control (RNA biosynthesis), protein degradation, nutrient uptake, various membrane channels (Solute transport), cell wall and photosynthesis are among some of the most responsive to up to seven stresses (Figure 3A). Conversely, more basal processes, such as protein biosynthesis/translocation, polyamine/amino acid/nucleotide metabolism, DNA damage response/chromatin organization/cell cycle, were markedly less stress-responsive (Figure 3A). These results indicate that *Marchantia* adjusts the physiology, energy, and nutrient uptake via various gene expression regulatory mechanisms to ensure the homeostasis of basic biological functions. Since younger gene families show a higher degree of organ-specific gene expression[40], we speculated that younger gene families could have evolved to provide adaptations to specialized environments and their abiotic stresses. However, we observed similar stress responsiveness patterns for the oldest gene families (Archaeplastida, Figure 3B, red line) and the much younger *Marchantia*-specific gene families (Mpo, green line), suggesting that younger gene families are not more active in stress responses, which is in line with our previous observation in *Cyanophora*[20].

Since plants are often exposed to stress combinations in nature and our understanding of how plants respond to these combinations is lacking[29,62], we have

exposed *Marchantia* to 19 possible combinations of the seven abiotic stresses (darkness+high light and cold+heat being mutually exclusive). Interestingly, purely physical stresses (cold, heat, darkness, and high light) showed a higher number of DEGs than chemical stresses (salt, mannitol, and nitrogen deficiency) (Figure 4A). We speculate that the physical stresses cause more DEGs because these stresses are able to permeate every cell and affect every protein (heat, darkness) and/or have a dramatic effect on the energy levels that have consequences on all processes (darkness, high light). Conversely, stresses such as mannitol and salt can be effectively contained by the action of ion transporters and osmolyte accumulation[63]. However, we did not observe any correlation between the number of DEGs and the effect on plant growth (Figure 4B); while salt and mannitol treatments caused the most dramatic growth defects, these two stresses also showed the lowest number of DEGs (Figure 4B). This observation is in line with our study in alga *Cyanophora*[20], suggesting that this phenomenon is likely to be found in other organisms.

We observed a complex interplay of the stress combinations, where, e.g., the combination of heat and prolonged darkness caused a unique upregulation of ~500 DEGs (Figure 4C, first column), but we also observed that certain stresses produce a core set of DEGs present in any stress combination (e.g., Figure 4C, fourth column of darkness combinations). Further analysis revealed that darkness-containing stresses show a similar pattern of DEGs, indicating that darkness is perceived as more severe stress by the plant and is thus able to dominate the DEG responses over other stresses (Figure 5D, E, dominance plots are red in darkness stresses). This dominance is caused by darkness-containing stresses having a similar pattern of downregulated biological pathways (Figure 5F, blue cells indicate downregulation). Taken together, the analysis allowed us to rank the strength of dominance of abiotic stresses as darkness (six stress combinations), heat and light (four combinations each), nitrogen deficiency, and cold (three each).

Interestingly, while salt and mannitol were the least dominant stresses and at the same time most similar in terms of DEGs (Figure 5D-E, SM is dark blue for similarity), when combined, they produced the most novel response (Figure 5D, E, dark blue for 'novel interaction' for SM, Figure 5F, SM is an outgroup). Other novel responses comprise cold+high light and cold+nitrogen deficiency (Figure 5F). The basis for the novel responses is unclear, but the unique DEG response found in the combination of cold and high light (Figure 5D, CL is dark blue; Figure 5F, CL is an outlier) and the enrichment of hydrogen peroxide removal pathway (Figure 5F) suggests that certain stress combinations can enhance one another. High light can lead to the production of oxidative stress by ROS, and plants regulate the photosynthetic ROS production by adjusting light-harvesting and electron consumption pathways[64]. However, since cold is known to slow down most metabolic reactions, the electron consumption pathways are likely unable to absorb the excess electrons, leading to ROS accumulation. Since ROS

can cause oxidative damage to a wide range of lipids, proteins, and nucleic acids[65] and is itself a signalling molecule (for recent reviews, see [66–68]), this resulted in the many novel response that we observed for cold+high light.

Since the diurnal cycle regulates a significant number of genes in most Archaeplastida [43], we set to augment our stress expression atlas with diurnal gene expression data. Similarly to other land plants and in contrast to algae [43], Marchantia's cell division is not synchronized by the diurnal cycle (Figure 6C). Interestingly, Marchantia seems to have two populations of genes that peak at the end of the day (phase 10) and the end of the night (phase 20, Figure 6B), which is also reflected in the expression patterns of numerous biological pathways (Figure 6D). We observed this previously in *Arabidopsis thaliana* and to lesser extent in rice, but not in *Physcomitrium patens, Selaginella moellendorffii* or *Picea abie*s, suggesting that this pattern is not dictated by the evolutionary distance. However, we observed that the diurnal gene expression is conserved between most members of Archaeplastida (Figure 6F), indicating that *Marchantia*, with its simplified regulatory network, is an excellent model to study diurnal gene expression.

To make our data more readily accessible, we provide an eFP browser for Marchantia, a popular tool allowing the visualization of gene expression by an 'electronic Fluorescent Pictograph'. Furthermore, we provide an updated CoNekT database with expression atlases of 13 species comprising various algae and land plants. The database provides tools to study gene expression, functional enrichment analyses of co-expression networks, and other comparative tools. These valuable tools will help further dissect the gene regulatory networks behind abiotic stress responses in *Marchantia* and other species Figure 7).

Through our analyses, we have demonstrated the similarity of *Marchantia* to other members of the Archaeplastida in terms of the responsiveness of gene families to stress, the dominance, and the novel response of *Marchantia* to combined stress, and the conserved diurnal expression of genes. We have validated the observations that when presented with more than one stress, plants have a more dominant response to the more severe stressor and that novel responses are adopted when plants are exposed to combined stress. This strengthens the call for more emphasis on the study of combined biotic and abiotic stresses in light of future challenges posed by climate change[33,62].

Regrettably, while our study suggests a lack of conservation in gene expression across stress and species, this might be convoluted by the difference in sample type (shoot, leaves, and whole organisms), the analysis of few representative species, and the sheer complexity of how plants respond to stress at the epigenetic (methylation of genes), transcriptomic (mRNA, microRNA, lncRNA) and proteomic (posttranslational modifications and activation) levels[69–73]. In addition, the broad level comparison in this study may not have been sufficient to sift out the similarities between basal and higher

plants due to the increased complexity of plants due to gene and genome duplications. Conservation of stress response in plants has been observed in other studies which looked at transcription factors and hormones[74–79]. Taking the various limitations and considerations outlined above, future studies of abiotic stress in plants should involve a multi-omics approach with the inclusion of more species for us to delve into the intricacies of the adaptations plants have developed over millions of years to respond to stress.

Acknowledgments

We want to thank the Nanyang Technological University start-up grant for funding. We want to thank the members of the Mutwil lab for their comments on the manuscript. We thank D. Maizels (http://www.scientific-art.com/) for the illustrations used for the eFP browser.

Author contributions

M.M. conceived the project, wrote the manuscript, Q.W.T. performed the single stress experiments, data analysis, updated the CoNekT database and wrote the manuscript, P.K.L performed the combined stress experiments, N.P. and A.P. provided the eFP browser, Z.C., provided the Marchantia cultures and feedback on growing Marchantia.

Other competing interests

The authors declare no competing interests.

References

1. Taylor, L. L., Banwart, S. A., Valdes, P. J., Leake, J. R. & Beerling, D. J. Evaluating the effects of terrestrial ecosystems, climate and carbon dioxide on weathering over geological time: a global-scale process-based approach. *Philos Trans R Soc Lond B Biol Sci* **367**, 565–582 (2012).

2. Wickett, N. J. *et al.* Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proc Natl Acad Sci U S A* **111**, E4859-4868 (2014).

3. Graham, L. E., Cook, M. E. & Busse, J. S. The origin of plants: body plan changes contributing to a major evolutionary radiation. *Proc Natl Acad Sci U S A* **97**, 4535–4540 (2000).

4. Bowman, J. L. Walkabout on the long branches of plant evolution. *Curr Opin Plant Biol* **16**, 70–77 (2013).

5. Rubinstein, C. V., Gerrienne, P., de la Puente, G. S., Astini, R. A. & Steemans, P. Early Middle Ordovician evidence for land plants in Argentina (eastern Gondwana). *New Phytol* **188**, 365–369 (2010).

6. Wellman, C. H., Osterloff, P. L. & Mohiuddin, U. Fragments of the earliest land plants. *Nature* **425**, 282–285 (2003).

7. Wheeler, J. A. Molecular Phylogenetic Reconstructions of the Marchantioid Liverwort Radiation. *bryo* **103**, 314–333 (2000).

8. Bowman, J. L. *et al.* Insights into Land Plant Evolution Garnered from the Marchantia polymorpha Genome. *Cell* **171**, 287-304.e15 (2017).

9. Ishizaki, K., Nishihama, R., Yamato, K. T. & Kohchi, T. Molecular Genetic Tools and Techniques for Marchantia polymorpha Research. *Plant Cell Physiol* **57**, 262–270 (2016).

10. Cesarino, I. *et al.* Plant science's next top models. *Ann Bot* **126**, 1–23 (2020).

11. Lampugnani, E. R. *et al.* Cellulose Synthesis – Central Components and Their Evolutionary Relationships. *Trends in Plant Science* **24**, 402–412 (2019).

12. Shimamura, M. Marchantia polymorpha: Taxonomy, Phylogeny and Morphology of a Model System. *Plant Cell Physiol* **57**, 230–256 (2016).

13. Eklund, D. M. *et al.* Auxin Produced by the Indole-3-Pyruvic Acid Pathway Regulates Development and Gemmae Dormancy in the Liverwort Marchantia polymorpha. *Plant Cell* **27**, 1650–1669 (2015).

14. Tanaka, D., Ishizaki, K., Kohchi, T. & Yamato, K. T. Cryopreservation of Gemmae from the Liverwort Marchantia polymorpha L. *Plant and Cell Physiology* **57**, 300–306 (2016).

15. Togawa, T. *et al.* Cryopreservation of Marchantia polymorpha spermatozoa. *J Plant Res* **131**, 1047–1054 (2018).

16. Borghi, M., Perez de Souza, L., Yoshida, T. & Fernie, A. R. Flowers and climate change: a metabolic perspective. *New Phytol* **224**, 1425–1441 (2019).

17. Bailey-Serres, J., Parker, J. E., Ainsworth, E. A., Oldroyd, G. E. D. & Schroeder, J. I. Genetic strategies for improving crop yields. *Nature* **575**, 109–118 (2019).

18. Walters, D., Walsh, D., Newton, A. & Lyon, G. Induced resistance for plant disease control: maximizing the efficacy of resistance elicitors. *Phytopathology* **95**, 1368–1373 (2005).

19. Abdelrahman, M., Jogaiah, S., Burritt, D. J. & Tran, L.-S. P. Legume genetic resources and transcriptome dynamics under abiotic stress conditions. *Plant, Cell & Environment* **41**, 1972–1983 (2018).

20. Ferrari, C. & Mutwil, M. Gene expression analysis of Cyanophora paradoxa reveals conserved abiotic stress responses between basal algae and flowering plants. *New Phytologist* (2019) doi:10.1111/nph.16257.

21. Zandalinas, S. I. *et al.* Systemic signaling during abiotic stress combination in plants. *PNAS* **117**, 13810–13820 (2020).

22. Thoen, M. P. M. *et al.* Genetic architecture of plant stress resistance: multi-trait genome-wide association mapping. *New Phytol* **213**, 1346–1362 (2017).

23. Cairns, J. E. *et al.* Identification of Drought, Heat, and Combined Drought and Heat Tolerant Donors in Maize. *Crop Science* **53**, 1335–1346 (2013).

24. Davila Olivas, N. H. *et al.* Genome-wide association analysis reveals distinct genetic architectures for single and combined stress responses in Arabidopsis thaliana. *New Phytologist* **213**, 838–851 (2017).

25. Nilson, S. E. & Assmann, S. M. The control of transpiration. Insights from Arabidopsis. *Plant Physiol* **143**, 19–27 (2007).

26. Sun, Z., Jin, X., Albert, R. & Assmann, S. M. Multi-level Modeling of Light-Induced Stomatal Opening Offers New Insights into Its Regulation by Drought. *PLOS Computational Biology* **10**, e1003930 (2014).

27. Zandalinas, S. I. *et al.* ABA is required for the accumulation of APX1 and MBF1c during a combination of water deficit and heat stress. *J Exp Bot* **67**, 5381–5390 (2016).

28. Zhou, R. *et al.* Screening and validation of tomato genotypes under heat stress using Fv/Fm to reveal the physiological mechanism of heat tolerance. *Environmental and Experimental Botany* **118**, 1–11 (2015).

29. Rivero, R. M., Mittler, R., Blumwald, E. & Zandalinas, S. I. Developing climate-resilient crops: improving plant tolerance to stress combination. *Plant J* (2021) doi:10.1111/tpj.15483.

30. Jiang, J. & Dehesh, K. Plastidial retrograde modulation of light and hormonal signaling: an odyssey. *New Phytologist* **230**, 931–937 (2021).

31. Martínez-Ruiz, A. *et al.* Specificity in S-Nitrosylation: A Short-Range Mechanism for NO Signaling? *Antioxidants & Redox Signaling* **19**, 1220–1235 (2013).

32. Stulemeijer, I. J. E. & Joosten, M. H. a. J. Post-translational modification of host proteins in pathogen-triggered defence signalling in plants. *Molecular Plant Pathology* **9**, 545–560 (2008).

33. Atkinson, N. J., Jain, R. & Urwin, P. E. The Response of Plants to Simultaneous Biotic and Abiotic Stress. in *Combined Stresses in Plants: Physiological, Molecular, and Biochemical Aspects* (ed. Mahalingam, R.) 181–201 (Springer International Publishing, 2015). doi:10.1007/978-3-319-07899-1_9.

34. Pandey, P., Ramegowda, V. & Senthil-Kumar, M. Shared and unique responses of plants to multiple individual stresses and stress combinations: physiological and molecular mechanisms. *Frontiers in Plant Science* **6**, 723 (2015).

35. Gamborg, O. L., Miller, R. A. & Ojima, K. Nutrient requirements of suspension cultures of soybean root cells. *Experimental Cell Research* **50**, 151–158 (1968).

36. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, (2014).

37. Schwacke, R. *et al.* MapMan4: A Refined Protein Classification and Annotation Framework Applicable to Multi-Omics Data Analysis. *Molecular Plant* **12**, 879–892 (2019).

38. Jin, J. *et al.* PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research* **45**, D1040–D1045 (2017).

39. Zheng, Y. *et al.* iTAK: A Program for Genome-wide Prediction and Classification of Plant Transcription Factors, Transcriptional Regulators, and Protein Kinases. *Mol Plant* **9**, 1667–1670 (2016).

40. Julca, I. *et al.* Comparative transcriptomic analysis reveals conserved programmes underpinning organogenesis and reproduction in land plants. *Nat Plants* **7**, 1143–1159 (2021).

41. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289–300 (1995).

42. Hughes, M. E., Hogenesch, J. B. & Kornacker, K. JTK-CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *Journal of Biological Rhythms* **25**, 372–380 (2010).

43. Ferrari, C. *et al.* Kingdom-wide comparison reveals the evolution of diurnal gene expression in Archaeplastida. *Nature Communications* **10**, (2019).

44. Rico-Reséndiz, F. *et al.* Transcriptional and Morpho-Physiological Responses of Marchantia polymorpha upon Phosphate Starvation. *International Journal of Molecular Sciences* **21**, 8354 (2020).

45. Buti, M. *et al.* Transcriptome profiling of short-term response to chilling stress in tolerant and sensitive Oryza sativa ssp. Japonica seedlings. *Funct Integr Genomics* **18**, 627–644 (2018).

46. Huang, L. *et al.* Comparative transcriptome sequencing of tolerant rice introgression line and its parents in response to drought stress. *BMC Genomics* **15**, 1026 (2014).

47. Domazet-Lošo, T., Brajković, J. & Tautz, D. A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends in Genetics* **23**, 533–539 (2007).

48. Das, P., Nutan, K. K., Singla-Pareek, S. L. & Pareek, A. Oxidative environment and redox homeostasis in plants: dissecting out significant contribution of major cellular organelles. *Frontiers in Environmental Science* **2**, 70 (2015).

49. Winter, D. *et al.* An ‘electronic fluorescent pictograph’ Browser for exploring and analyzing large-scale biological data sets. *PLoS ONE* **2**, (2007).

50. Proost, S. & Mutwil, M. CoNekT: An open-source framework for comparative genomic and transcriptomic network analyses. *Nucleic Acids Research* **46**, W133–W140 (2018).

51. Vogt, T. Phenylpropanoid Biosynthesis. *Molecular Plant* **3**, 2–20 (2010).

52. Bais, H. P., Vepachedu, R., Gilroy, S., Callaway, R. M. & Vivanco, J. M. Allelopathy and Exotic Plant Invasion: From Molecules and Genes to Species Interactions. *Science* **301**, 1377–1380 (2003).

53. Gräfe, K. & Schmitt, L. The ABC transporter G subfamily in Arabidopsis thaliana. *Journal of Experimental Botany* **72**, 92–106 (2021).

54. Nakayama, T. *et al.* Specificity analysis and mechanism of aurone synthesis catalyzed by aureusidin synthase, a polyphenol oxidase homolog responsible for flower coloration. *FEBS Lett* **499**, 107–111 (2001).

55. Pégeot, H. *et al.* The poplar Phi class glutathione transferase: expression, activity and structure of GSTF1. *Frontiers in Plant Science* **5**, 712 (2014).

56. Ferrari, C. *et al.* Expression Atlas of Selaginella moellendorffii Provides Insights into the Evolution of Vasculature, Secondary Metabolism, and Roots. *Plant Cell* **32**, 853–870 (2020).

57. Bokszczanin, K. *et al.* Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. *Frontiers in Plant Science* **4**, 315 (2013).

58. You, J. & Chan, Z. ROS Regulation During Abiotic Stress Responses in Crop Plants. *Front Plant Sci* **6**, 1092 (2015).

59. Sanchez, D. H. *et al.* Comparative functional genomics of salt stress in related model and cultivated plants identifies and overcomes limitations to translational genomics. *PLoS ONE* **6**, (2011).

60. Van Leeuwen, H. *et al.* Natural variation among Arabidopsis thaliana accessions for transcriptome response to exogenous salicylic acid. *Plant Cell* **19**, 2099–2110 (2007).

61. Wang, F. *et al.* Comparative transcriptomics reveals differential gene expression related to Colletotrichum gloeosporioides resistance in the octoploid strawberry. *Frontiers in Plant Science* **8**, (2017).

62. Mahalingam, R. Consideration of Combined Stress: A Crucial Paradigm for Improving Multiple Stress Tolerance in Plants. in *Combined Stresses in Plants: Physiological, Molecular, and Biochemical Aspects* (ed. Mahalingam, R.) 1–25 (Springer International Publishing, 2015). doi:10.1007/978-3-319-07899-1_1.

63. Hanin, M., Ebel, C., Ngom, M., Laplaze, L. & Masmoudi, K. New Insights on Plant Salt Tolerance Mechanisms and Their Potential Use for Breeding. *Frontiers in Plant Science* **7**, 1787 (2016).

64. Lima-Melo, Y. *et al.* Photoinhibition of Photosystem I Provides Oxidative Protection During Imbalanced Photosynthetic Electron Transport in Arabidopsis thaliana. *Frontiers in Plant Science* **10**, 916 (2019).

65. McCord, J. M. The evolution of free radicals and oxidative stress. *Am J Med* **108**, 652–659 (2000).

66. Mittler, R. ROS Are Good. *Trends in Plant Science* **22**, 11–19 (2017).

67. Czarnocka, W. & Karpiński, S. Friend or foe? Reactive oxygen species production, scavenging and signaling in plant response to environmental stresses. *Free Radical Biology and Medicine* **122**, 4–20 (2018).

68. Mullineaux, P. M., Exposito-Rodriguez, M., Laissue, P. P. & Smirnoff, N. ROS-dependent signalling pathways in plants and algae exposed to high light: Comparisons with other eukaryotes. *Free Radical Biology and Medicine* **122**, 52–64 (2018).

69. Chang, Y.-N. *et al.* Epigenetic regulation in plant abiotic stress responses. *Journal of Integrative Plant Biology* **62**, 563–580 (2020).

70. Mazzucotelli, E. *et al.* Abiotic stress response in plants: When post-transcriptional and post-translational regulations control transcription. *Plant Science* **174**, 420–431 (2008).

71. Franke, K. R. *et al.* Analysis of Brachypodium miRNA targets: evidence for diverse control during stress and conservation in bioenergy crops. *BMC Genomics* **19**, 547 (2018).

72. Barrero-Gil, J. & Salinas, J. Post-translational regulation of cold acclimation response. *Plant Science* **205–206**, 48–54 (2013).

73. Schütze, K., Harter, K. & Chaban, C. Post-translational regulation of plant bZIP factors. *Trends in Plant Science* **13**, 247–255 (2008).

74. Castro, P. H. *et al.* Phylogenetic analysis of F-bZIP transcription factors indicates conservation of the zinc deficiency response across land plants. *Sci Rep* **7**, 3806 (2017).

75. Wu, T.-Y. *et al.* Evolutionarily conserved hierarchical gene regulatory networks for plant salt stress response. *Nat. Plants* **7**, 787–799 (2021).

76. Balderas-Hernández, V. E., Alvarado-Rodríguez, M. & Fraire-Velázquez, S. Conserved versatile master regulators in signalling pathways in response to stress in plants. *AoB Plants* **5**, plt033 (2013).

77. Chen, W. J. & Zhu, T. Networks of transcription factors with roles in environmental stress response. *Trends in Plant Science* **9**, 591–596 (2004).

78. Ju, C. *et al.* Conservation of ethylene as a plant hormone over 450 million years of evolution. *Nature Plants* **1**, 1–7 (2015).

79. Lieberman-Lazarovich, M., Yahav, C., Israeli, A. & Efroni, I. Deep Conservation of cis-Element Variants Regulating Plant Hormonal Responses. *The Plant Cell* **31**, 2559–2572 (2019).

Supplementary figures

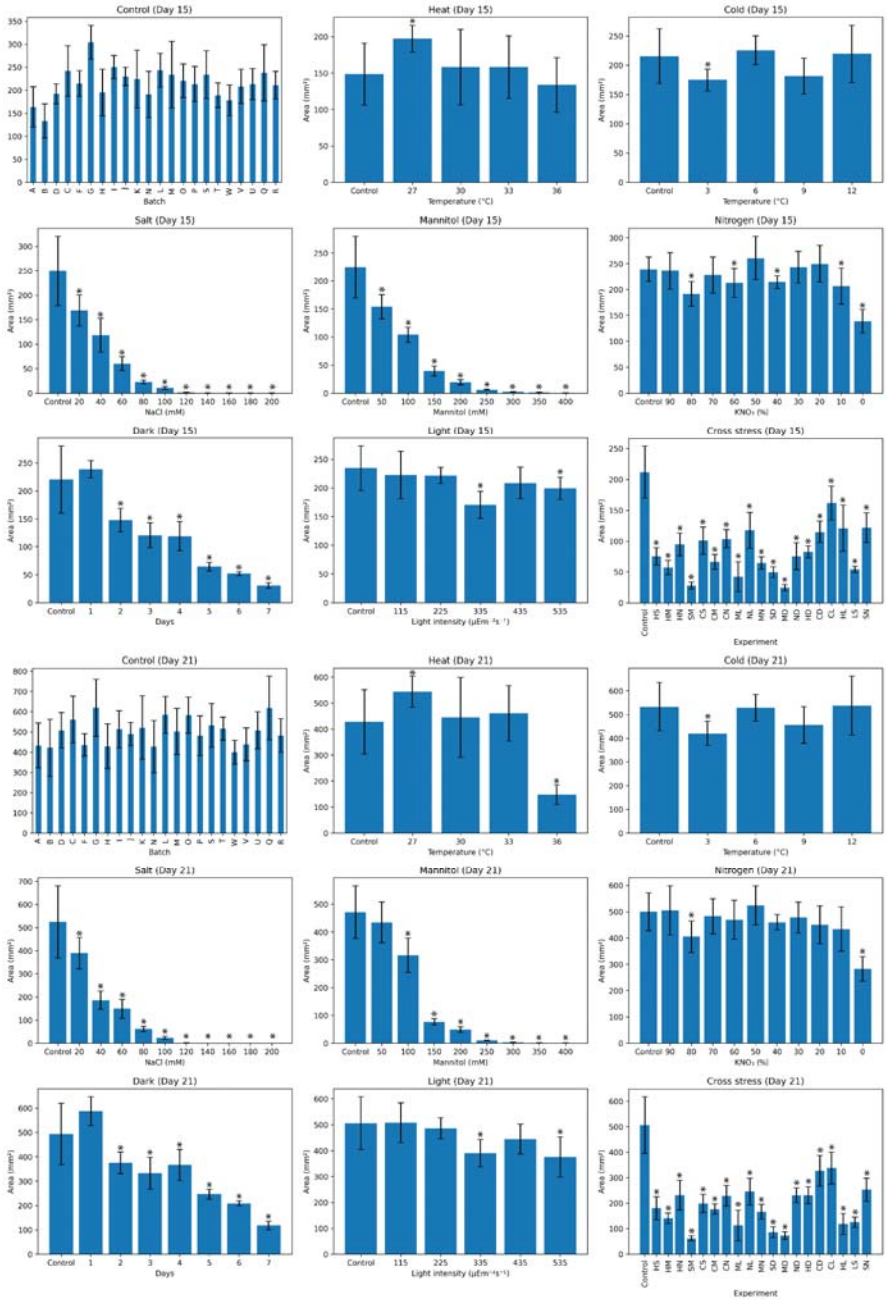


Figure S1. Area of plants at 15 days (day of harvest) and day 21 (6 days post-harvest) for the single and combined stresses. Each bar represents measurements from at least 4 plants. Error bars are represented by standard deviation and significance was determined by Student's two-tailed, t-test, $p < 0.05$.

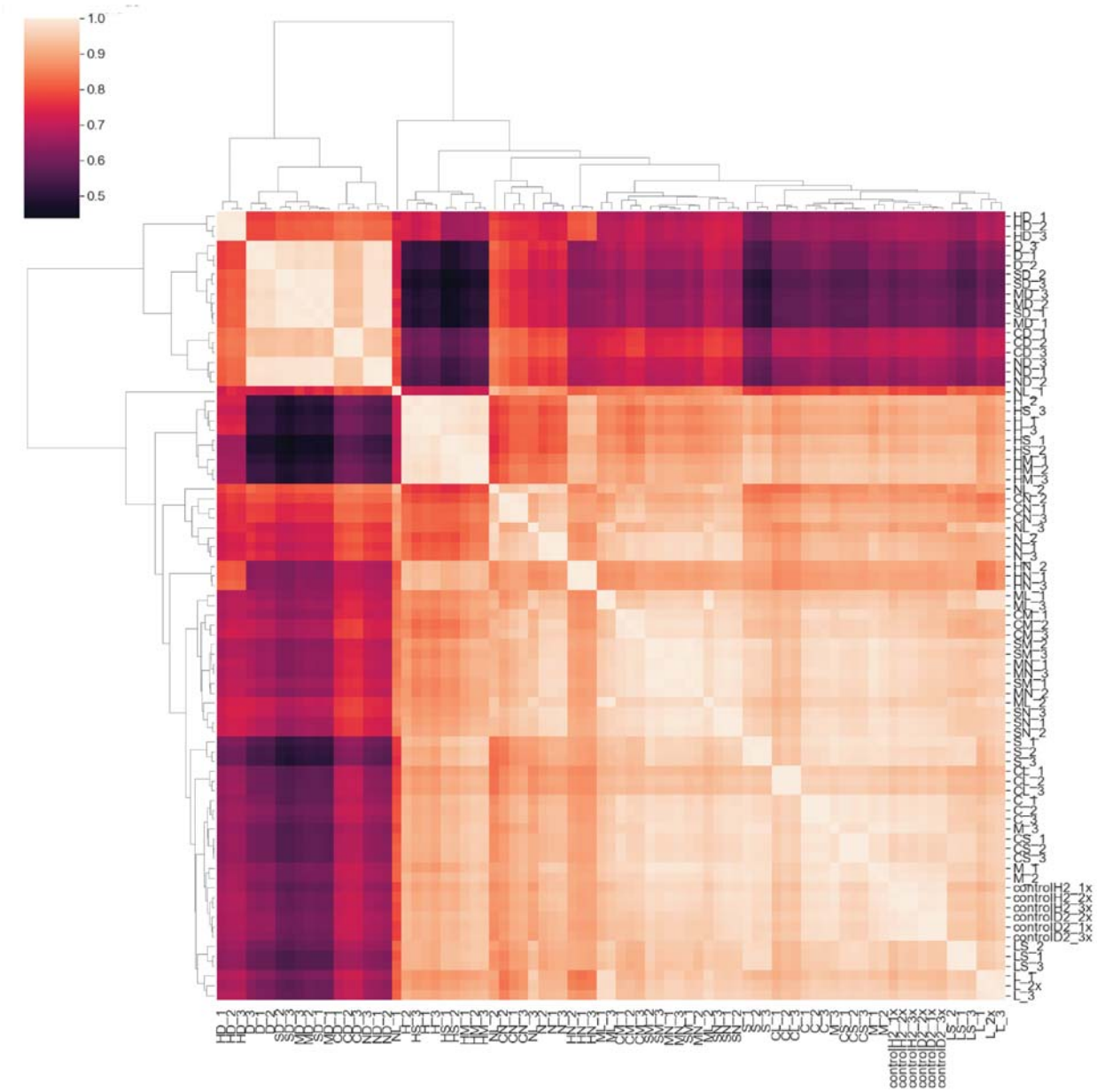


Figure S2. Hierarchical clustering of the Marchantia stress experiments. Transcript per million (TPM) gene expression values were scaled with standard scaler and clustered with seaborn.clustermap() function.

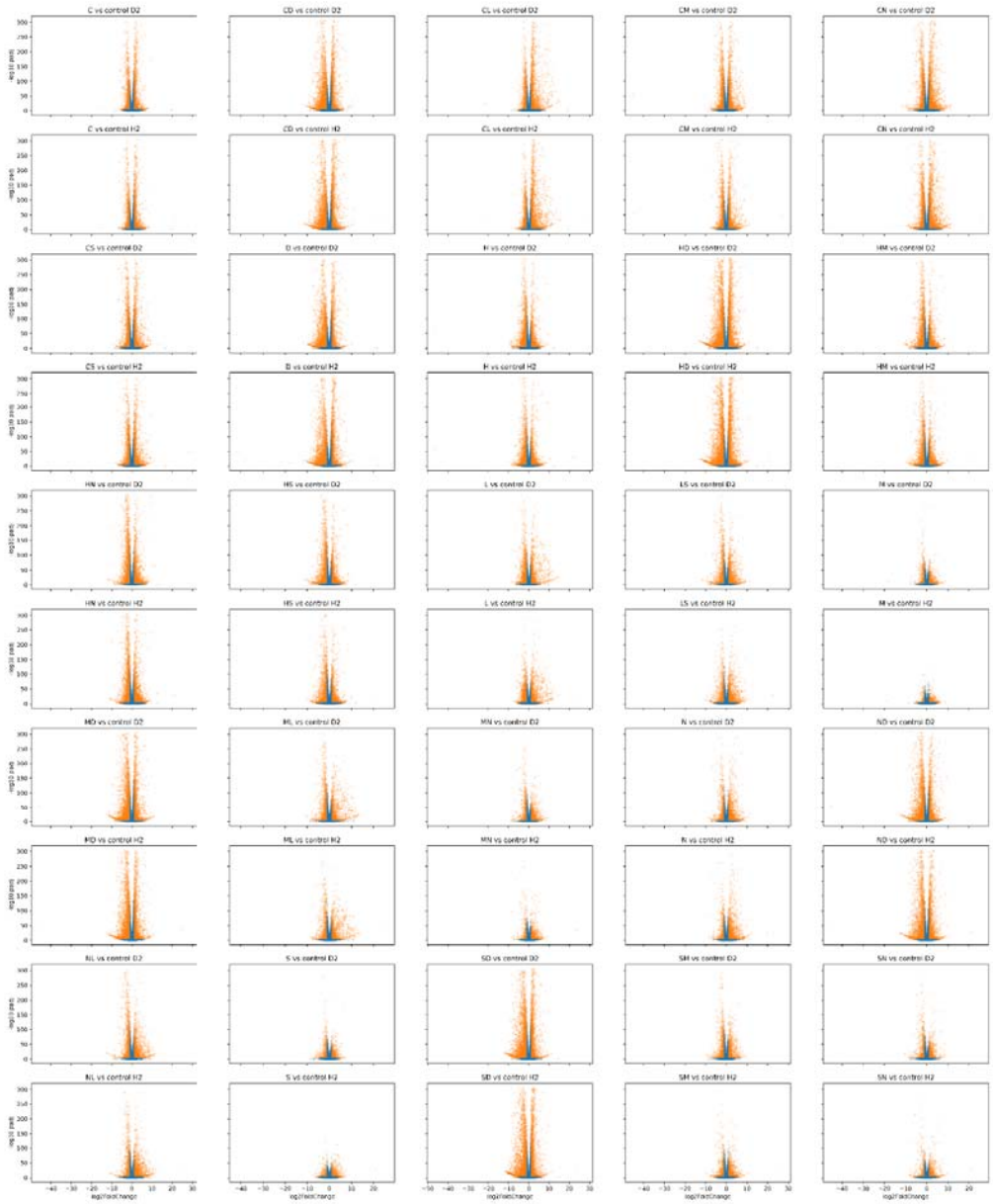


Figure S3. Volcano plots of stress experiments. The x-axis represents $\log_2$-fold change, the y-axis indicates the $-\log_{10}$ of adjusted p-value. Each point represents a gene that has an absolute $\log_2$-fold change more than 1 and is significantly (adjusted p-value<0.05, orange color) or not significantly (p-value>0.05, blue point) differentially expressed.

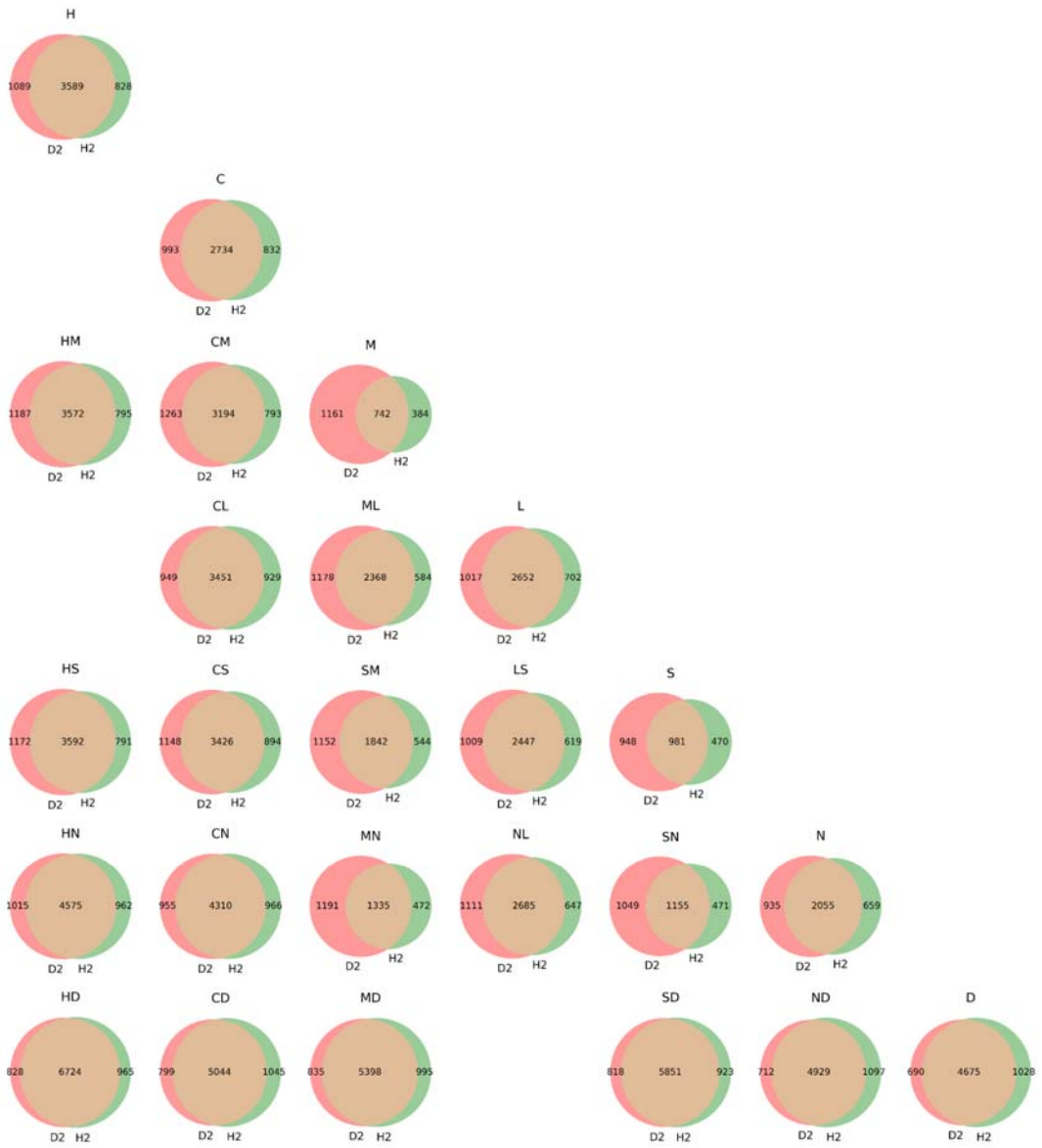


Figure S4. Comparison of found differentially expressed genes for the single and combined stressed. For each stress combination, we estimated significant DEGs (adjusted p-value<0.05) by using control D2 (red circle) and control H2 (green circle). The orange areas indicate the overlap between the found DEGs of the two controls. The overlapping DEGs were used for further study. The following abbreviations are used for the description of stresses: cold (C), dark (D), heat (H), light (L), mannitol (M), nitrogen deficiency (N) and salt (S). Combination of stresses is indicated by two letters, where e.g., HD indicates heat+darkness treatment.

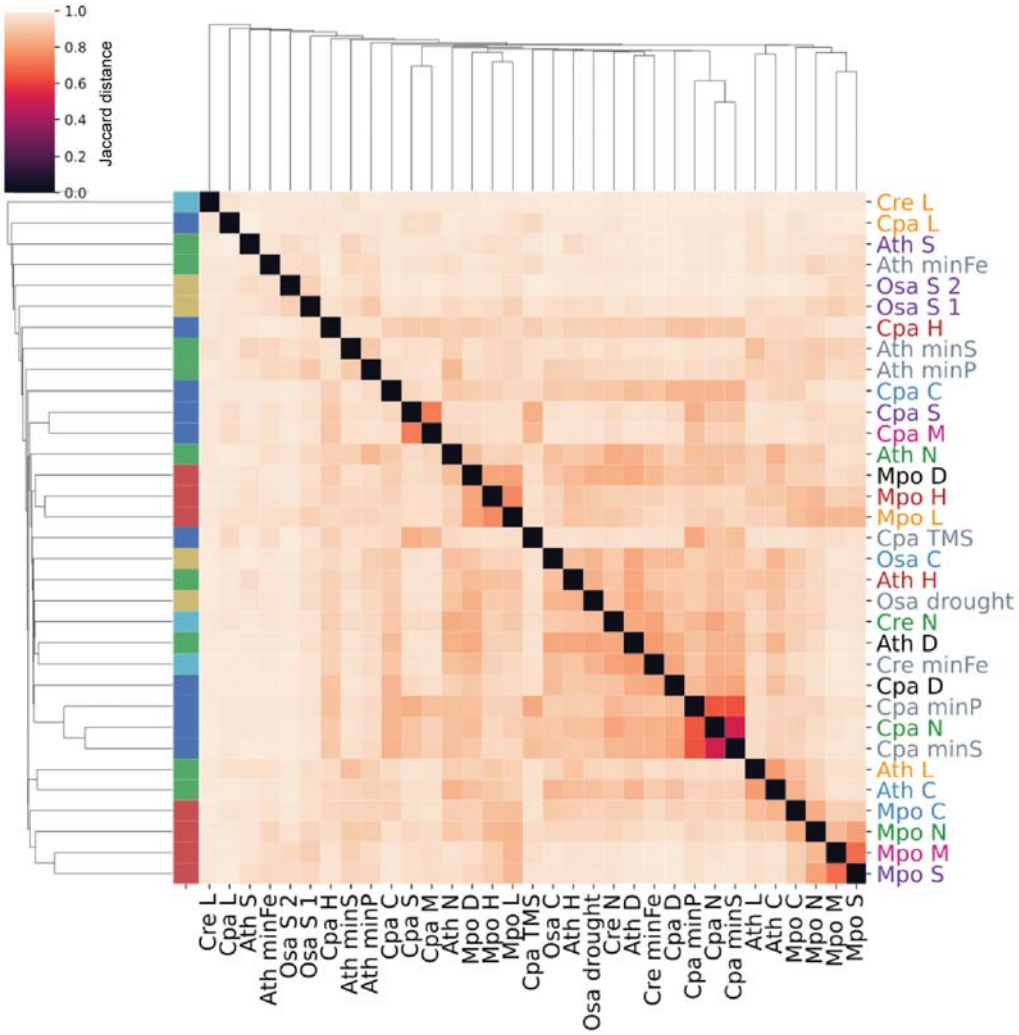


Figure S5. Clustering of stress/species similarities by using differentially expressed gene families. Gene families were classified according to the conservation of differential gene expression and similarities between the differential gene families were calculated by jaccard index. The color of each cell indicates the Jaccard Distance between any two stress/species. Species are colour-coded blue, light blue, red, yellow and green on the left column for *C. paradoxa, C. reinhardtii, M. polymorpha, O. sativa and A. thaliana* respectively while labels for heat, cold, salt, mannitol, light, dark, nitrogen and other stresses are coloured in red, blue, purple, magenta, yellow, black, green and grey respectively. The following abbreviations are used for the description of stresses: cold (C), dark (D), heat (H), light (L), mannitol (M), nitrogen deficiency (N), salt (S), iron deficiency (minFe), sulfur deficiency (minS) and phosphorus deficiency (minP).

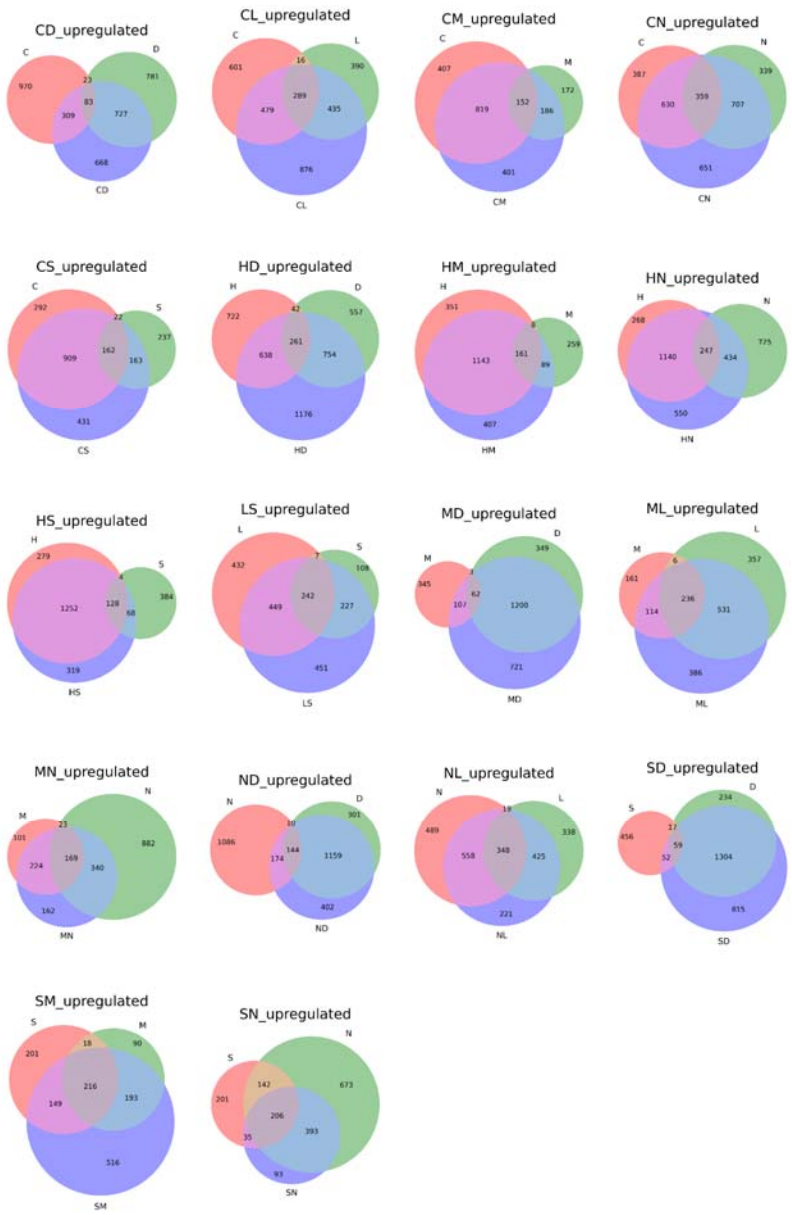


Figure S6. Venn diagrams of upregulated genes in single and combined stresses. The abbreviates are cold (C), dark (D), heat (H), light (L), mannitol (M), nitrogen deficiency (N) and salt (S). Combination of stresses is indicated by two letters and purple circles. The first and second stress in a pair is colored red and green (e.g., cold is red in CD), while the combined stress is colored purple. The sizes of the circles and intersections and the numbers within indicate the number of significant DEGs.

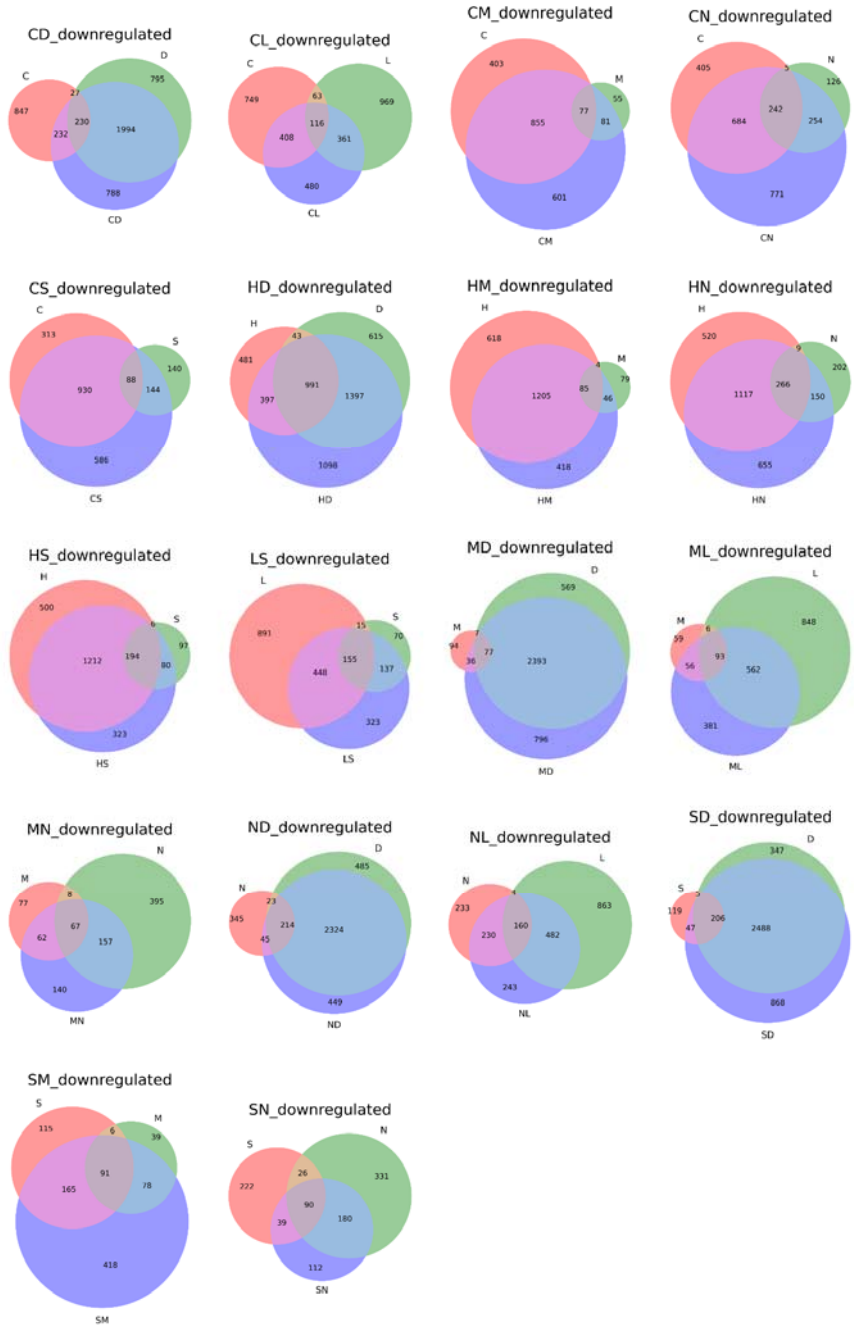


Figure S7. Venn diagrams of downregulated genes in single and combined stresses. The abbreviates are cold (C), dark (D), heat (H), light (L), mannitol (M), nitrogen deficiency (N) and salt (S). Combination of stresses is indicated by two letters and purple circles. The first and second stress in a pair is colored red and green (e.g., cold is red in CD), while the combined stress is colored purple. The sizes of the circles and intersections and the numbers within indicate the number of significant DEGs.

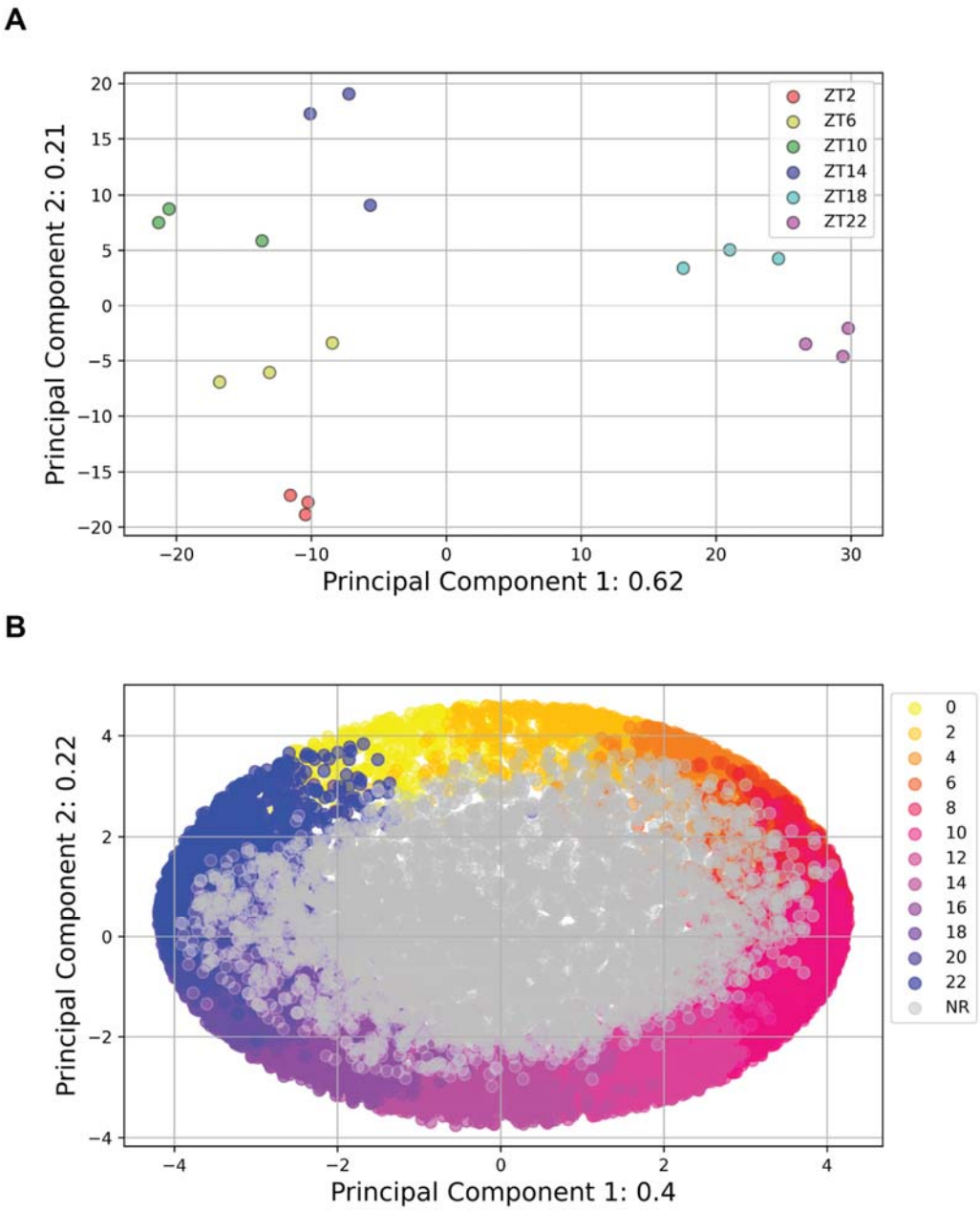


Figure S8. Analysis of diurnal data. A) Principal Component Analysis (PCA) of the diurnal gene expression data. Each point represents one RNA-seq sample. The different zeitgeber (ZT) time points are color coded. The data was scaled with standard scaler. B) PCA analysis of genes. The genes were color-coded by their lag value. Gray genes are not rhythmic.

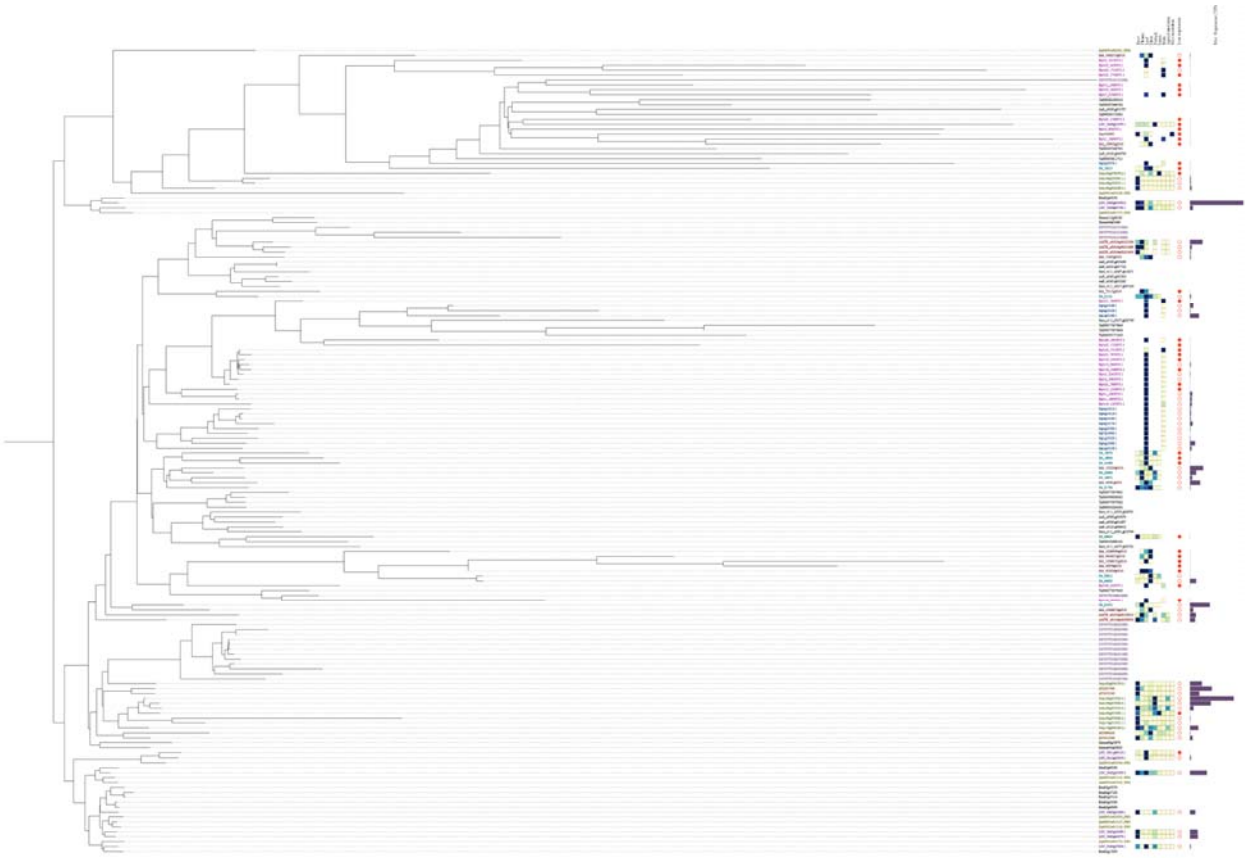


Figure S9. Phylogenetic Tree of PAL OG_02_0000119 (Land Plants).

Supplementary Tables

Supplementary Table 1. Summary of experiment label to experiment conditions and mapping statistics from kallisto.

Supplementary Table 2. TPM normalised gene expression matrix for *Marchantia* experiments

Supplementary Table 3. Raw gene expression matrix for *Marchantia* experiments

Supplementary Table 4. PCC values within *Marchantia* experiment replicates

Supplementary Table 5. Details on source, sample and stress conditions of other species in interspecies comparison

Supplementary Table 6. Summary of DESeq2 output of *Marchantia* stress experiments

Supplementary Table 7. List of genes that are identified to be differentially expressed in both controls with corresponding mapman bins and annotation. 'UP' and 'DOWN' in L2FC_D2 and L2FC_H2 columns indicate significant up and downregulation when compared against control D2 and H2 respectively.

Supplementary Table 8. Summary of DESeq2 output of *Oryza sativa* stress experiments

Supplementary Table 9. List of genes that are identified to be differentially expressed in *Oryza sativa*.

Supplementary Table 10. Annotation of transcription factors for Cpa, Cre, Mpo, Osa and Ath

Supplementary Table 11. Annotation of kinases for Cpa, Cre, Mpo, Osa and Ath

Supplementary Table 12. Status of orthogroups in corresponding species and stress condition. 'UP', 'DOWN', 'AMB', 'NC' and 'None' indicates that all differentially expressed genes are upregulated, downregulated, up and downregulated, no differentially expressed genes and no genes in orthogroup respectively.

Supplementary Table 13. Summary of age classification of orthogroups, combined status of response to stress in various organisms for non-species specific orthogroups and response to stress in species-specific orthogroups.

Supplementary Table 14. Orthogroup file used for orthogroup comparisons, obtained from Julca et al., 2021

Supplementary Table 15. JTK_Cycle output of diurnal gene expression for *Marchantia*

Supplementary Table 16. Cold-Nitrogen deficiency specific genes as obtained from conekt.plant.tools