

**Oil body formation in *Marchantia polymorpha* is controlled by
MpC1HDZ and serves as a defense against arthropod herbivores**

Facundo Romani, ^{1,7} Elizabeta Banic, ^{2,7} Stevie N. Florent, ³ Takehiko Kanazawa, ⁴ Jason Q.D. Goodger, ⁵ Remco Mentink, ² Tom Dierschke, ^{3,6} Sabine Zachgo, ⁶ Takashi Ueda, ⁴ John L. Bowman, ³ Miltos Tsiantis, ^{2*} Javier E. Moreno ^{1*}

1. Instituto de Agrobiotecnología del Litoral, Universidad Nacional del Litoral – CONICET, Facultad de Bioquímica y Ciencias Biológicas, Centro Científico Tecnológico CONICET Santa Fe, Colectora Ruta Nacional No. 168 km. 0, Paraje El Pozo, Santa Fe 3000, Argentina.

2. Department of Comparative Development and Genetics, Max Planck Institute for Plant Breeding Research, Carl-von-Linne-Weg 10, 50829, Cologne, Germany.

3. School of Biological Sciences, Monash University, Melbourne, Victoria 3800, Australia.

4. Division of Cellular Dynamics, National Institute for Basic Biology, Nishigonaka 38, Myodaiji, Okazaki, Aichi, 444-8585, Japan

5. School of BioSciences, The University of Melbourne, Parkville, Victoria 3010, Australia.

6. Botany Department, School of Biology and Chemistry, Osnabrück University, Barbarasträße 11, 49076 Osnabrück, Germany.

7. Equal contribution

*Correspondence: tsiantis@mpipz.mpg.de ; javier.moreno@santafe-conicet.gov.ar

KEYWORDS

MpC1HDZ, HD-ZIP, Marchantia, oil bodies, oil body cell, cell differentiation, biotic defense, sesquiterpenes, terpene synthases, transcription factor, evolution, liverwort.

SUMMARY

The origin of a terrestrial flora in the Ordovician required adaptation to novel biotic and abiotic stressors. Oil bodies, a synapomorphy of liverworts, accumulate secondary metabolites, but their function and development are poorly understood. Oil bodies of *Marchantia polymorpha* develop within specialized cells as one single large organelle. Here, we show that a CLASS I HOMEODOMAIN LEUCINE-ZIPPER (C1HDZ) transcription factor controls the differentiation of oil body cells in two different ecotypes of the liverwort *M. polymorpha*, a model genetic system for early divergent land plants. In flowering plants, these transcription factors primarily modulate responses to abiotic stress including drought. However, loss-of-function alleles of the single ortholog gene, MpC1HDZ, in *M. polymorpha* did not exhibit phenotypes associated with abiotic stress. Rather Mpc1hdz mutant plants were more susceptible to herbivory and total plant extracts of the mutant exhibited reduced antibacterial activity. Transcriptomic analysis of the mutant revealed a reduction in expression of genes related to secondary metabolism that was accompanied by a specific depletion of oil body terpenoid compounds. Through time lapse imaging we observed that MpC1HDZ expression maxima precede oil body formation indicating that MpC1HDZ mediates differentiation of oil body cells. Our results indicate that *M. polymorpha* oil bodies, and MpC1HDZ, are critical for defense against herbivory but not for abiotic stress-tolerance. Thus, C1HDZ genes were co-opted to regulate separate responses to biotic and abiotic stressors in two distinct land plant lineages.

INTRODUCTION

The emergence of land flora in the Ordovician and the subsequent radiation of plants into terrestrial environments were major evolutionary events in the history of Earth. This radical breakthrough for life on the planet entailed the evolution of both new developmental programs and biochemical pathways [1]. Early land plants were faced with novel biotic and abiotic stressors, including a desiccating aerial environment, increased solar radiation, and greater interspecific competition, including herbivory. This transition was accompanied by a dramatic increase in the diversity of secondary metabolites and the complexity of transcriptional regulation, to protect plants from herbivores and pathogens, and also abiotic stressors such as desiccation and light intensity [2-4].

Liverworts represent one of the earliest diverging land plant lineages, with a predicted Ordovician-Silurian origin of all three major extant liverwort clades [5, 6]. One synapomorphy of this lineage is the presence of oil bodies, specialized organelles containing an array of lipophilic secondary compounds [7]. In some liverworts, multiple oil bodies are found in nearly all differentiated cells of both gametophyte and sporophyte generations. Whilst in other liverworts, such as *Marchantia polymorpha* (*M. polymorpha*), oil bodies are large and differentiate in only a subset of isolated cells referred to as idioblasts. They originate in young meristematic cells in a relatively constant proportion [8]. It has long been known that the liverwort fragrance is associated to the oil body presence [9], and that species-specific differences in taste correlate with oil body chemical composition [10]. Early biochemical analyses identified numerous terpenoids as major chemical constituents of liverwort oil bodies [11, 12], and later studies expanded this list to include diterpenes, bis(bibenzyls), and additional compounds [13-16]. Pfeffer, who gave oil bodies their name (*Ölkörper*), noted they were membrane bound [17], and later studies confirmed a single membrane [18, 19] in which isoprenoid biosynthesis enzymes are localized [16].

Oil body function remains contested, with suggested defensive roles against both abiotic and biotic stressors. Roles protecting against a wide variety of abiotic stressors have been proposed, including

cold and osmotic stress [7], desiccation [20-22], high light intensity [23], UV radiation [24, 25], and metabolic stress [18]. In addition, a role against herbivory was postulated by Stahl, who presented snails with fresh or alcohol-leached liverworts and noted that snails preferred the latter and did not consume the former to any extent [26]. An anti-herbivory role is also supported by fossilized evidence of herbivore oil body avoidance in Devonian liverworts [27]. Furthermore, the chemical constituents of oil bodies lead to growth inhibition of bacterial and fungal pathogens [28-32]. Despite a large body of literature describing the presence and diversity of oil body secondary metabolites [14, 15, 33, 34], less is known of the regulatory circuits leading to oil body cell differentiation.

Synthesis of specialized metabolites in vascular plants is tightly regulated by transcription factors (TFs), with some orthologs having conserved and others novel roles in liverworts [32, 35-38]. Class I HOMEODOMAIN LEUCINE-ZIPPER (C1HDZ) TFs genes are conserved across streptophytes with a single homolog predicted to have been present in the common-ancestor of land plants [39]. In angiosperms, C1HDZ genes are well-known regulators of tolerance to abiotic stress [40, 41], but also control other molecular functions including leaf, root and silique development and vascular patterning [42-44]. However, little is known about the physiological role of C1HDZ in bryophytes [45]. *M. polymorpha* provides a model to investigate their functional evolution since the genome encodes only one single C1HDZ ortholog (MpC1HDZ) [2]. Here, we investigate MpC1HDZ functionality and report that it exhibits a divergent function relative to its angiosperms orthologs.

RESULTS

MpC1HDZ controls the development of oil body cells

MpC1HDZ, the solitary C1HDZ *M. polymorpha* ortholog, possesses all conserved protein motifs [39] and predicted structural features previously defined for C1HDZ proteins (Figure 1A, Figure S1A). MpC1HDZ exhibits nuclear localization in transient expression assays in tobacco plants (Figure

S1B). Using two different genetic backgrounds (WT = Australian, BoGa = Botanical Garden Osnabrück, German accession), we targeted gRNAs to the HDZ domain generating mutant alleles with premature STOP codons, deletion of the whole HDZ domain or deletion of the entire *locus* (Figure 1A). Four lines were clonally purified through gemmae generations: *Mpc1hdz-1^{ge}* (male) and *Mpc1hdz-2^{ge}* (female) harbor frameshift mutations that generate early stop codons and were generated in the Australian accession. *Mpc1hdz-3^{ge}* and *Mpc1hdz-4^{ge}* were generated in the BoGa accession deleting the whole CDS or the HDZ domain, respectively (Figure 1A; nomenclature as in [46]).

In all cases, mutant plants did not show gross phenotypic defects and completed their life cycle in a similar manner and timing as wild type (Figure S2). To understand the physiological role of *MpC1HDZ*, we performed a detailed phenotypic characterization of the mutant alleles. The most conspicuous phenotype of *Mpc1hdz* mutants is a striking reduction in the number of oil bodies, both in gemmae and thalli, in two different genetic backgrounds (Figure 1B,C). *M. polymorpha* oil body cells are observed as brown dots or highly-reflective cells, with 20-50 oil bodies per gemmae, mostly located marginally (Figure 1B,C). Whereas mutant gemmae lacked oil bodies, in the thallus and in the parenchymatic region we found a small number of oil bodies in the tissue (Figure S3). It is known that specific environmental conditions, such as starvation and field non-axenic cultivation, affect oil body cell number [15]. Wild type plants grown in Petri-dishes developed fewer oil body cells compared to non-axenic grown plants and starvation-treated plants (Figure S3). In contrast, the number of oil body cells of the *Mpc1hdz* mutant remained low in both conditions (Figure S3).

To examine oil body distribution throughout the life cycle, we crossed *Mpc1hdz-2^{ge}* with *Mpc1hdz-1^{ge}*. The resultant sporophyte produced spores of the two genotypes and sporelings were phenotypically wild type. After 18-days, oil bodies were visible in wild type sporelings, but were totally absent in *Mpc1hdz* sporelings (Figure S3C). Thus, we conclude that *MpC1HDZ* is a positive regulator of oil body formation across all organs tested and developmental stages.

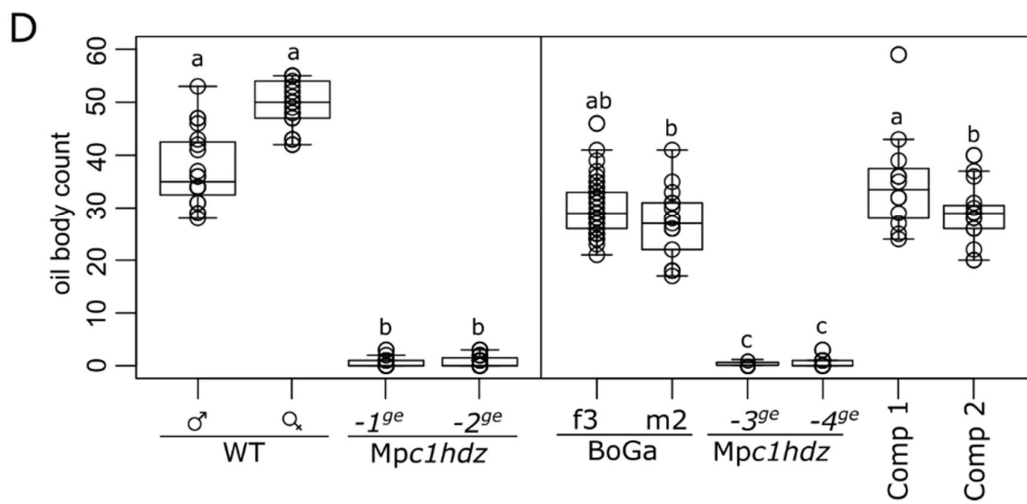
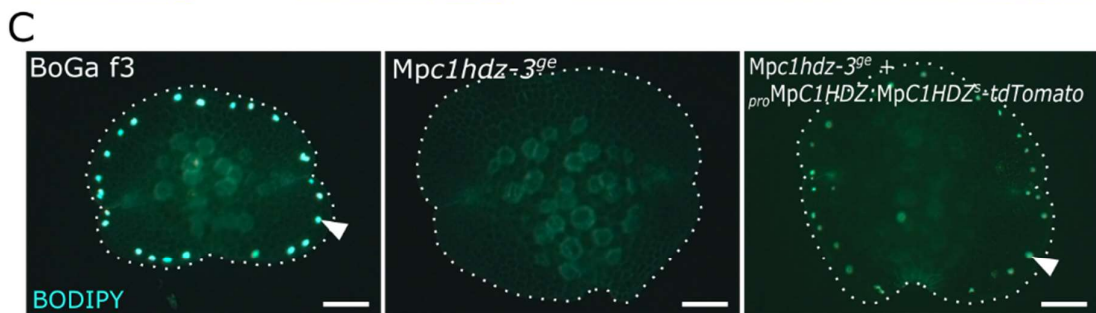
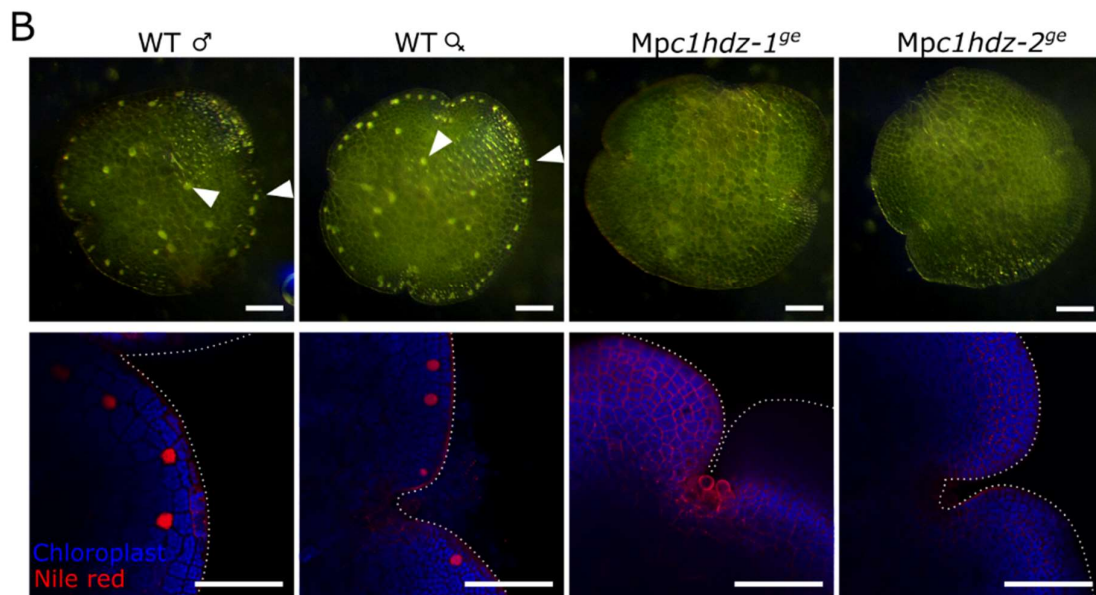
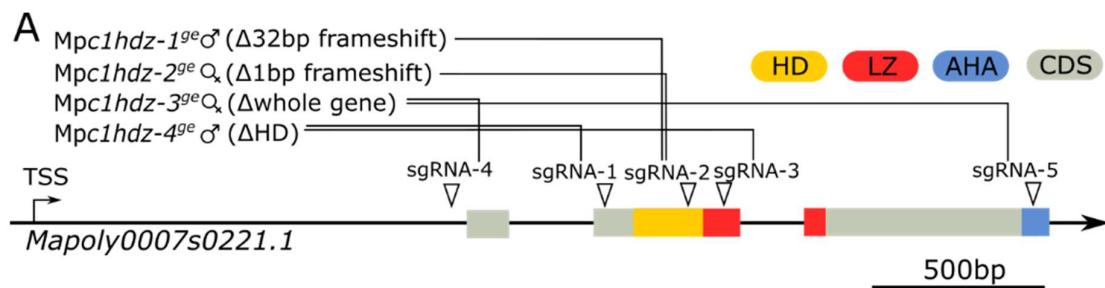


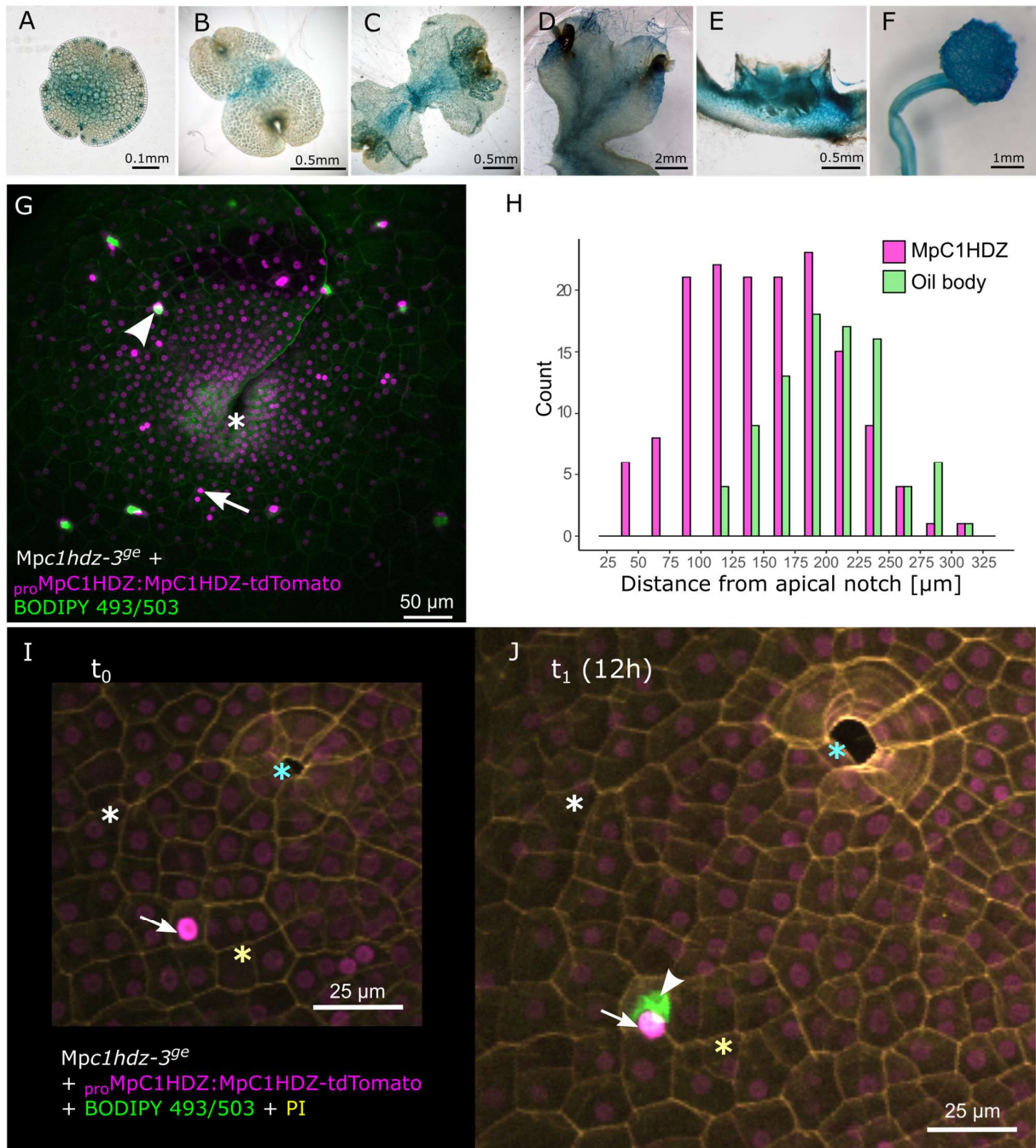
Figure 1. Oil body number is strongly reduced in *Mpc1hdz* plants. (A) Schematic representation of the *MpCIHDZ* gene domains and location of designed CRISPR gRNAs targeting the genomic locus (arrowheads). Colored boxes represent the three conserved protein domains. The combination of gRNAs used to generate frameshifts deletions in each case is shown. (B) Phenotype of freshly collected gemmae from wild-type and *Mpc1hdz* mutant plants under light microscopy (upper panel); confocal microscopy of Red Nile stained gemmae (bottom, panel). Red Nile fluorescence is shown in red and chloroplast auto-fluorescence in blue. (C) Oil body formation is restored in gemmae expressing *MpCIHDZ* translational reporter *proMpCIHDZ^{3kb}:MpCIHDZ-tdTomato* and visualized by BODIPY staining. (D) Number of oil bodies per gemma in wild-type, mutant and complemented plants in both genetic backgrounds were quantified and represented as boxplot (n>12). Comp 1: *Mpc1hdz-3^{ge}* + *proMpCIHDZ^{3kb}:MpCIHDZ-Citrine*; Comp 2: *Mpc1hdz-3^{ge}* + *proMpCIHDZ^{3kb}:MpCIHDZ-tdTomato*. Statistical difference was tested using ANOVA followed by Tukey HSD *p* value < 0.01; letters indicate statistically significant groups.

To confirm the role of *MpCIHDZ* in the oil body phenotype, we complemented the mutant using the endogenous promoter of *MpCIHDZ* driving *MpCIHDZ-tdTomato* (*proMpCIHDZ^{3kb}:MpCIHDZ-tdTomato*) in the BoGa ecotype, which rescued the oil body phenotype (Figure 1C-D). In contrast, when *MpCIHDZ* expression was driven by either of two constitutive promoters (*proEF1:MpCIHDZ*, *pro35S:MpCIHDZ*) no viable plants were obtained (Figure S4), suggesting that constitutive overexpression is lethal (data not shown).

The *MpCIHDZ* promoter is active in thallus and gemmalings with high expression in oil body cells

To better understand the relationship between *MpCIHDZ* and oil body cells, we studied the gene expression pattern with transcriptional reporters (*proMpCIHDZ:GUS*) using the putative regulatory region of each genetic background. In 3-4 day-old gemmalings, strong GUS activity was observed in oil body cells (Figure 2A). At later stages, GUS activity was detected in the center of the thallus and in meristematic regions (Figure 2B,C), at the notch extending into the midrib, including the gemma

151 cups (Figure 2C-E), and broadly in reproductive organs (Figure 2F). A similar transcriptional reporter
 152 activity was also observed in the BoGa ecotype (Figure S5).



154 **Figure 2. Strong *MpC1HDZ* expression precedes oil body formation.** (A-F) Representative
 155 pictures of *proMpC1HDZ^{4.6kb}::GUS* staining (from left to right) of fresh gemmae and 1-week-old plant
 156 (Australian ecotype) light microscopy, stereoscopic microscopy of transverse sections of 3-week-old
 157 gemma cup grown in non-axenic conditions. (G) A representative apical notch area of *Mpc1hdz-3*

gemmalings complemented with *proMpCIHDZ^{3kb}:MpCIHDZ-tdTomato* and stained with BODIPY 493/503. Sum slices projection of a confocal stack shows *proMpCIHDZ^{3kb}:MpCIHDZ-tdTomato* signal in magenta and BODIPY in green. Asterisk marks the apical notch, arrowhead marks an oil body, and arrowhead marks an exceptionally bright nucleus. (H) Distribution of MpCIHDZ-tdTomato expression maxima and oil body distances from apical notch in 3 day-old gemmalings, quantified from 2D projections as in G. Quantification method is described in Figure S7. All bright nuclei above a defined MpCIHDZ-tdTomato fluorescence intensity threshold and all oil bodies from 11 replicates were counted. Significance was tested using Wilcoxon rank sum test, *p* value < 0.05. (I-J) *Mpc1hdz-3* gemmalings complemented with *proMpCIHDZ^{3kb}:MpCIHDZ-tdTomato* (magenta), stained with PI (yellow) and BODIPY 493/503 (green) and observed using CLSM microscopy in 12 h intervals (Standard deviation z-stack projections). Asterisks of the same color mark the corresponding landmarks in different time-points. Arrow points to nucleus of the cell that developed an oil body (arrowhead) within the next 12 h. All 13 observed transition events are shown in Figure S8 and the event shown here corresponds to number 6 (S8 I, J). Scale bars are shown in the picture.

To further test the role of MpCIHDZ in responses to environmental conditions, we used the transcriptional reporter in a wide range of stressors (Figure S6). Unlike in angiosperms [40], MpCIHDZ was not induced by abiotic stress treatments or ABA (Figure S6). We also tested biotic stress related treatments, such as salicylic acid (SA), wounding, or bacterial elicitors, which did not change the expression activity of MpCIHDZ promoter (Figure S6). However, among environmental conditions known to induce oil bodies, we found a strong induction of MpCIHDZ expression in the meristematic region when plants were grown in non-axenic conditions (soil pots), but not in starvation (Figure S6).

MpCIHDZ expression maxima precede oil body formation

To understand the dynamics of oil body formation and MpCIHDZ expression, we used a translational reporter *proMpCIHDZ^{3kb}:MpCIHDZ-tdTomato* that can complement the mutant phenotype. As observed in tobacco transient assays, fluorescently labelled MpCIHDZ-tdTomato was also nuclear

localized in stable *M. polymorpha* transgenic lines (Figure 2G, Figure S1). In 3 day-old gemmalings, high-expressing cells were located close to the notch in undifferentiated meristematic cells (Figure 2G). MpC1HDZ-tdTomato expression levels dropped gradually from the apical notch to distant cells, except for some isolated cells where MpC1HDZ-tdTomato accumulation became markedly higher compared to neighboring cells (Figure 2G). Using confocal microscopy to quantify the fluorescence of MpC1HDZ-tdTomato and oil bodies stained using BODIPY dye, we found an association between MpC1HDZ-tdTomato accumulation and oil body cells (Figure 2H). Closer to the apical notch sporadic cells strongly expressed MpC1HDZ-tdTomato without BODIPY signal in their vicinity. However, in cells more distant from the apical notch, the brighter MpC1HDZ-tdTomato nuclear signal was associated with BODIPY-stained oil bodies (Figure 2G,H). This initial correlation was further confirmed with cell tracking experiments. We observed that 11/13 newly detectable oil bodies in the field of view formed in cells that had exceptionally high MpC1HDZ-tdTomato expression in previous time point (Figure 2I,J, Figure S8). These results suggest that MpC1HDZ expression maxima precede oil body formation. Taken together, our results indicate that oil body cell differentiation is positively regulated by the MpC1HDZ.

Mpc1hdz mutants show normal abiotic stress responses

Ectopic expression of C1HDZ genes from different angiosperms in Arabidopsis enhances tolerance to water stress such as water deprivation, osmotic stress and salt stress [40, 47-51]. Moreover, some Arabidopsis loss-of-function C1HDZ alleles showed a reduced sensitivity to abscisic acid (ABA) treatments, consistent with a role in water stress tolerance [52, 53].

Using established protocols to manipulate abiotic stress conditions, we applied different treatments including water deprivation, osmotic and salt stress, to quantify the impact on plant growth, *i.e.*: area growth, thallus biomass. Our results showed no altered growth responses of the mutant lines compared to the wild type (Figure 3A-C; Figure S9). We also performed a water-loss experiment of

211 detached thalli and an ABA-dose response germination assay. We did not find a differential response
 212 of the mutant in either case (Figure 3D-E). To test whether *MpC1HDZ* expression is regulated by
 213 abiotic stress, we performed RT-qPCR analysis under NaCl and ABA treatments at two time-points,
 214 including also other members of the same protein family. Unlike the strong induction observed for
 215 *MpDHNI* [54], *MpHDZ* expression remained unaltered or were slightly reduced under these
 216 experimental conditions (Figure S10).

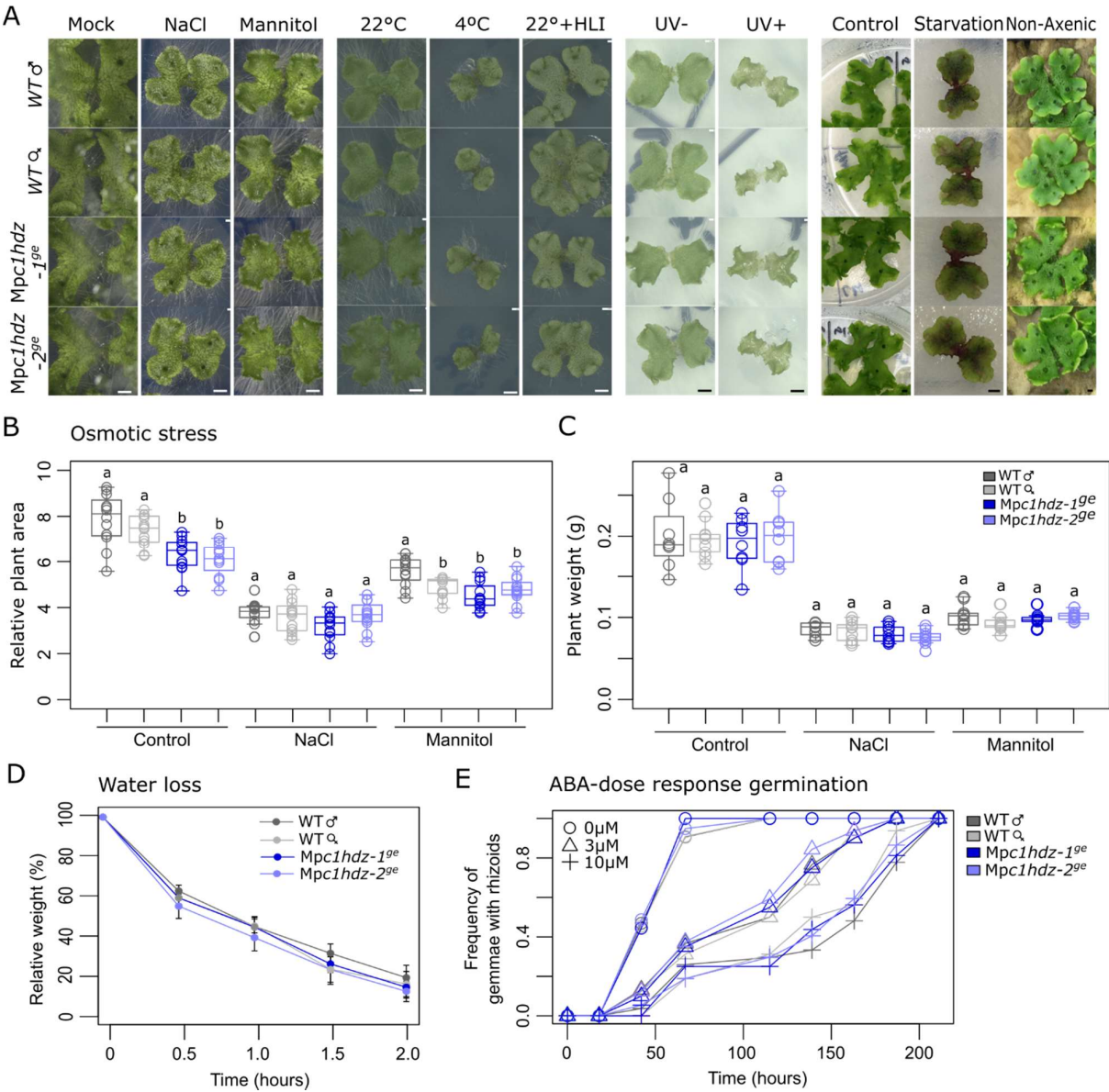


Figure 3. *Mpc1hdz* plants have an intact growth response to abiotic stress. (A) Representative picture of phenotypic effect of abiotic stress treatments on *Mpc1hdz* and wild-type plants as explained in methods section. Treatments are indicated at the top of illustrative pictures. Scale bars (2mm) are indicated in the lower panels. Individual values are indicated as circles. (B-D) Quantification of exposed area (B) and weight (C) of 3-week-old plants under osmotic stress treatments. Individual values are indicated as circles. Full quantifications for each treatment are shown in Figure S7. (D) Water loss experiments were done using 3-week-old wild-type and *Mpc1hdz* plants placed on filter paper towel and weighed every 30 min for 2 h at room temperature. Bars show the average of three biological replicates $\pm$ SE of the mean. No statistical difference was observed using ANOVA followed by Tukey HSD after 1 hour of treatment (p .value > 0.05). (E) Gemmae dormancy under ABA treatments at different concentrations and time points (left, $n \sim 25$). No statistical difference was found testing equal frequencies against wild-type using Exact Binomial Test (p .value > 0.05) at 115 h.

Oil bodies have long been proposed to protect liverworts against other abiotic stresses [55-57]. Therefore, we examined the response of *Mpc1hdz* plants to a wide range of other abiotic stressors. We grew plants in axenic control conditions for one week and subsequently transferred plants to different stress environments and measured growth parameters under cold stress, high light irradiance and UV light treatments (Figure S9). We also measured growth parameters in plants subjected to treatments that promote the formation of oil bodies (Figure S9). We did not observe significant differences in growth parameters among wild type and *Mpc1hdz* plants in any of the tested conditions. Overall, our results suggest that neither oil bodies nor *MpC1HDZ* are involved in the abiotic stress responses of *M. polymorpha* that *C1HDZ* gene function observed in angiosperms is not conserved throughout land plants.

***Mpc1hdz* mutants has an impaired biotic defense response**

Since oil bodies are rich sesquiterpenes stores and these compounds are known to play a central role in the ecology of plant-biotic interactions, together with old experiments associating the presence of oil bodies to the feeding behavior of snails [26], we investigated if oil bodies play a role in the biotic

response of *M. polymorpha*. We performed a choice experiment using wild type and *Mpc1hdz* plants to the terrestrial isopod *Armadillidium vulgare* (pill bug; Crustacea). In this case, we monitored the total thalli area consumed by starved pill bugs for 24 h. The total area lost from mutant alleles was significantly larger compared to respective wild-type plants in both genetic ecotypes (Figure 4A-C, Figure S11). Thus, isopods prefer mutant plants over wild-type, consistent with the possible role of oil bodies as a defense against arthropods herbivory.

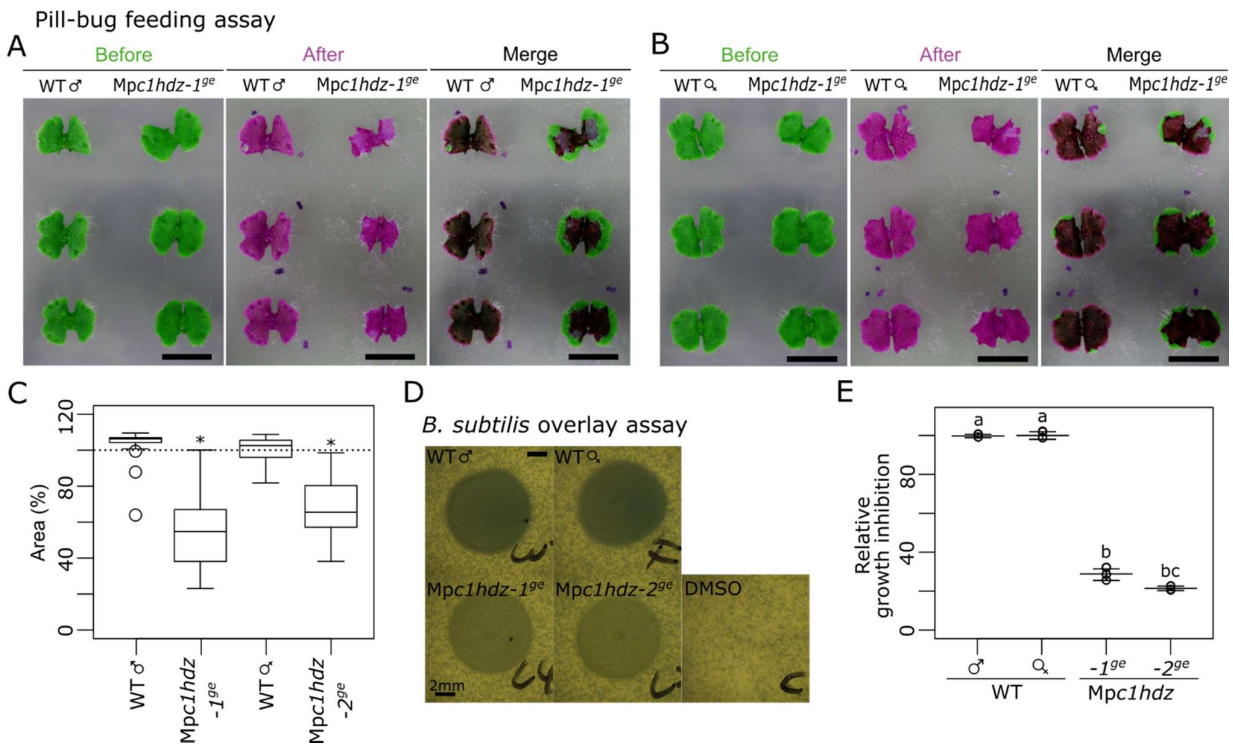


Figure 4. Phenotypic analysis of *Mpc1hdz* plants challenged with biotic stressors. Feeding assay was performed using starved pill bugs (*Armadillidium vulgare*) and 10-day-old thalli were co-cultivated for 24 hours. In order to contrast thallus areas, plants were colored in green (before) and magenta (after). Male wild-type and *Mpc1hdz-1^{ge}* thalli (A). Female wild-type and *Mpc1hdz-2^{ge}* thalli (B). Scale bars represent 1 cm. (C) Thallus damage quantification was performed using ImageJ and data represented as box plots. Thallus area change indicates the ratio of thallus area after and before (n = 36). Statistical differences are marked as asterisks and computed pairwise *t*-test (*p* value < 0.01) comparing each mutant with their respective control. (D) *Bacillus subtilis* overlay assay using ethanolic lipid extracts from wild-type and *Mpc1hdz* mutant thalli dissolved in DMSO.

Representative images of experimental results (three replicates were performed with consistent results). Scale bars = 2mm. (E) Quantification of relative growth inhibition as density of colonies relative to wild-type. The graph shows the mean with individual values ($n = 3$). Statistical difference was tested using ANOVA followed by Tukey HSD p .value < 0.01 ; letters indicate statistically significant groups.

As several compounds produced in oil bodies also have antibiotic activity [32, 58], they may provide further protection against biotic assault. To test this, we applied ethanolic extracts dissolved in DMSO from wild-type and *Mpc1hdz* thalli to the Gram (+) soil bacteria, *Bacillus subtilis*, using an overlay assay. We observed that antibacterial activity was reduced in *Mpc1hdz* extracts in a dose-dependent manner, around one order of magnitude compared to wild-type extracts (Figure 4D,E, Figure S11). Thus, in contrast to abiotic stresses, the striking phenotypic effects observed when challenged with biotic stressors provide convincing evidence that oil bodies are critical for biotic defense responses.

MpC1HDZ alters the expression of terpenoid-related genes

To better understand the molecular mechanisms underlying the altered defense response of the *Mpc1hdz* mutant, we compared the genome-wide transcript profiles of wild-type and *Mpc1hdz* plants grown in control (axenic) and non-axenic conditions. Here, we show the results of consistently differentially expressed genes between both RNA-seq experiments (Figure 5A,B). In control conditions, we identified 217 down-regulated genes ($\log_2(\text{FC}) < -1$, adj. p .value < 0.05) and 131 up-regulated genes ($\log_2(\text{FC}) > 1$, adj. p value < 0.05) in *Mpc1hdz* mutant plants (Figure 5, Table S1). Whereas, in non-axenic conditions, we identified 602 genes down-regulated and 267 up-regulated (Figure 5, Table S2). Additionally, we performed enrichment analysis over differentially expressed gene groups using an in-house pipeline of automatically annotated gene lists including GO-terms, biological processes and protein families (Table S3). Among down-regulated genes, there was an

enrichment in biological processes associated with secondary metabolism. Moreover, we found enrichment in proteins associated with terpene synthases (TPS, Table S3). Among, up-regulated genes, we did not find any significant enrichment.

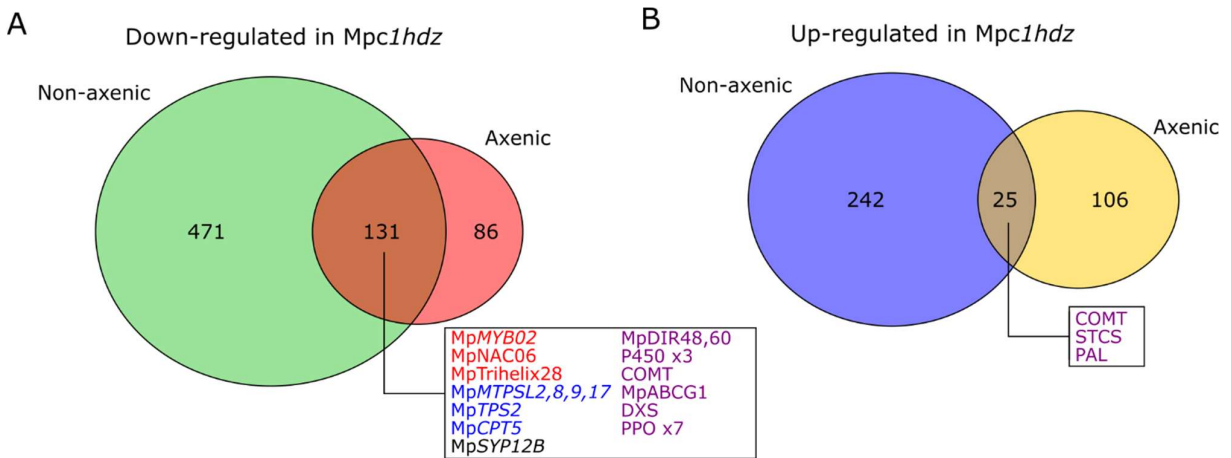


Figure 5. Transcriptomic analysis of *Mpc1hdz* mutants. Venn diagrams of group of genes down-regulated (A, $\log_2(\text{FC}) < -1$, $\text{adj.}p.\text{value} < 0.05$) and up-regulated (B, $\log_2(\text{FC}) > 1$, $\text{adj.}p.\text{value} < 0.05$) in 3-week-old vegetative thalli of *Mpc1hdz* compared to wild type in control (axenic) and non-axenic conditions as described in Methods section. Manually annotated genes associated with TFs (red), terpenoid-related biosynthesis (blue) and other secondary metabolite biosynthetic processes are highlighted. The full data is available at Table S1 and S2.

Within down-regulated genes associated with secondary metabolism (Figure 5A) and down-regulated in both experiments, we found five *MpTPS* (Figure 5A), including mono- and di- TPS with a bacterial, fungal or plant origin [33, 59]. We also detected other genes involved in the biosynthesis of terpene precursors, including a deoxyxylulose 5-phosphate synthase (*Mp4g14720*, *MpDXS*) and a cis-prenyltransferase (*Mp3g18510*, *MpCPT5*) gene (Figure 5A). In addition, three TF genes (*Mp3g07510*, *MpMYB02*; *Mp5g01530*, *MpNAC06* and *Mp2g16340*, *MpTrihelix28*) are down-regulated (Figure 5A). *MpMYB02* is known to regulate cyclic bis(bibenzyl) acids biosynthesis [35] which accumulates in oil bodies [15]. The mutant also exhibited significantly reduced expression of

MpSYPI2B (*Mp4g20670*), a N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) protein with specific localization at the oil body membrane of *M. polymorpha* [60]. Finally, we also identified other down-regulated genes that might be associated with the secondary metabolism, including a catechol-O-methyltransferase (*Mp5g06850*, MpCOMT), two dirigent proteins (Mp3g24140, MpDIR48 and Mp3g24210, MpDIR60), an ABC transporter (*Mp8g13070*, MpABCG1), and cytochrome P450 coding genes.

Oil body specific terpenoids are depleted in Mpc1hdz mutant plants

The strong reduction of oil bodies and the low expression of some *TPS* genes and other genes related to secondary metabolism, led us to explore the terpenoid profile of Mpc1hdz vegetative thalli. With the aim of evaluating the impact of environmental conditions on oil body chemistry, we also compared the terpenoid profile of plants grown under control, non-axenic, and starvation conditions (Figure S13A-C). In agreement with previous results [15], GC-FID and GC-MS analyses of thalli hexane extracts showed that non-axenic growth and starvation significantly increased the amount of terpenoid-related compounds compared to wild-type backgrounds, with starvation the most inductive condition (Figure S13D). Interestingly, male backgrounds showed higher levels of terpenoids than female ones, even in the mutant alleles (Table S4).

More importantly, the chemical profile of Mpc1hdz thalli revealed a specific decrease of terpenoid-related compounds (**2-7**), such as cis-thujopsene, β -chamigrene, β -himachalene and 5-hydroxy-gurjunene (Figure 6A,B). Most compounds that accumulated under oil body inducing conditions in wild-type plants were concomitantly reduced in mutant plants, suggesting they are located in oil bodies as suggested previously [15, 16] (Figure 6A,B). The most abundant monoterpene d-limonene (**1**) and the unidentified terpenoid U13 (**10**) (Figure 6 C,E; [33]) were strongly depleted in the mutants but they accumulated to similar levels in control and non-axenic conditions (Fig 6C, Figure S14). In

332 contrast, the most abundant diterpene, phytol (**11**) [33] (Figure 6D), fatty acid compounds (**8,9,12,13**)
333 (Figure 6E) and phytosterols, also present in plant extracts, did not significantly change in the mutant
334 or among growth conditions (Table S4). Taken together, these results support that oil body-localized
335 terpenoid biosynthesis is affected in *Mpc1hdz* mutants.

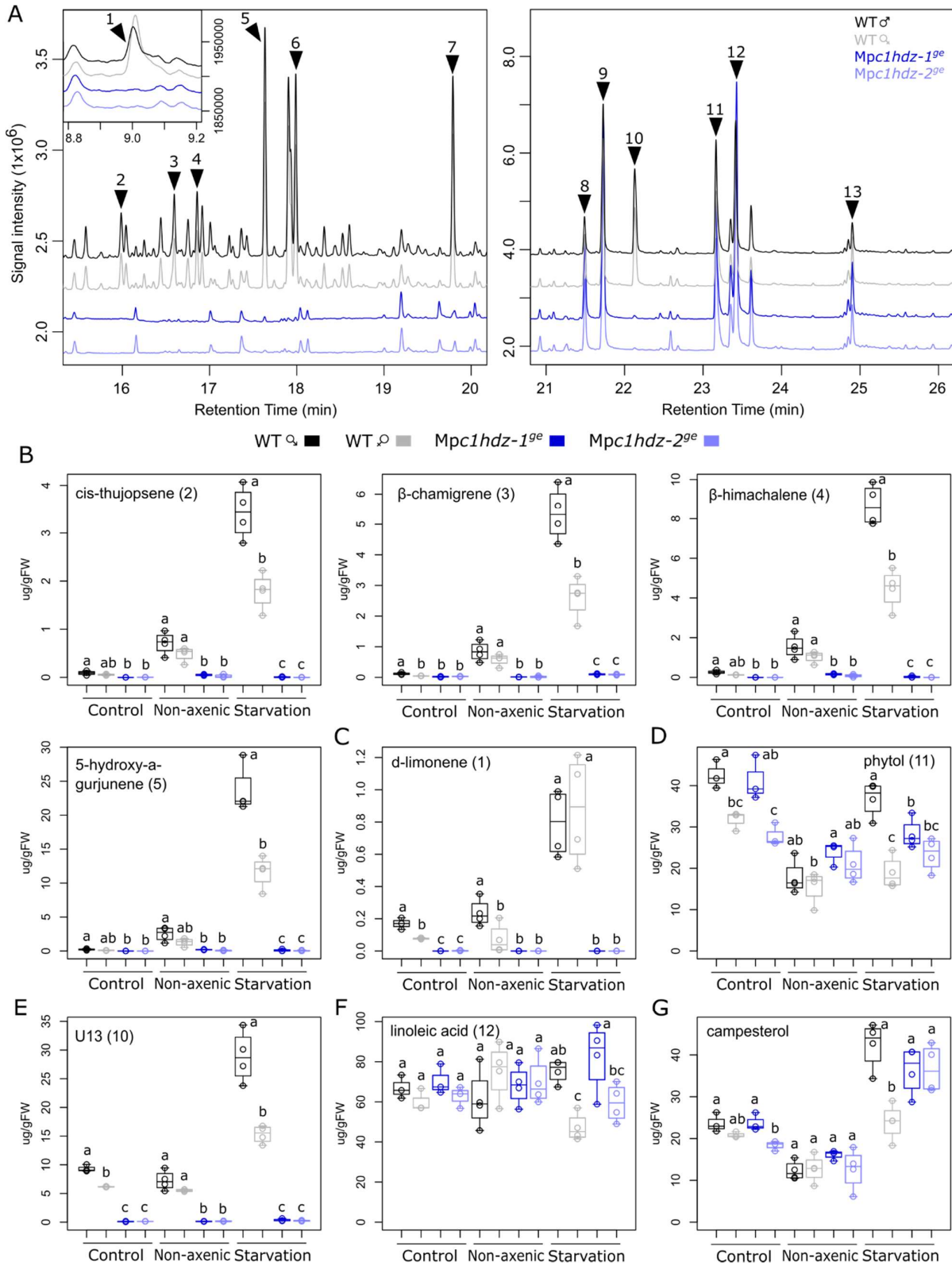


Figure 6. GC-FID analysis of *Mpc1hdz* mutants. (A) Representative GC-FID chromatogram obtained from hexane extracts under starvation treatment of *Mpc1hdz* and wild-type plants. Three

different ranges are shown. Arrow points represents metabolites: limonene (1), cis-thujopsene (2), β -chamigrene (3), β -himachalene (4), 5-hydroxy-gurjunene (5), viridiflorol (6), cyclocolorenone (7), stearic acid methyl ester (8), palmitic acid (9), unidentified terpenoid U13 (10), phytol (11), linoleic acid (12), timnodonic acid (13). Quantification of representative major metabolites in *Mpc1hdz* and wild-type plants under three different treatment as described before (control, non-axenic, and starvation) and represented as boxplots (n = 3-4): total sesquiterpenes (B), monoterpenes (C), diterpenes (D) and fatty acids (E). Statistical difference among genotypes was tested using ANOVA followed by Tukey HSD *p*.value < 0.05 for each treatment; letters indicate statistically significant groups.

DISCUSSION

C1HDZ genes have been previously characterized for their roles during abiotic stress response in angiosperms, more specifically to water stress tolerance in *Arabidopsis*, coffee, cotton, rice and sunflower [40]. Likely due to functional redundancy in angiosperms, the majority of C1HDZ mutants in angiosperms exhibit subtle phenotypes [39], with some exceptions. Homologs of *LATE MERISTEM IDENTITY 1/REDUCED COMPLEXITY* alter leaf morphology via local growth repression at early developmental stages [61, 62]. Given that *M. polymorpha* possesses a single C1HDZ gene, this is the first time that a full loss-of-function mutant is generated for this subfamily. Interestingly, the lack of a functional MpC1HDZ did not generate major pleiotropic effects as observed for other TF families in *M. polymorpha*. Using a wide array of experimental set-ups, we found no evidence supporting a physiological role for MpC1HDZ in the abiotic stress response of *M. polymorpha* (Figure 3, Figure S9). Further, in the moss *Physcomitrella patens*, only 1 out of 17 C1HDZ genes responded to dehydration and mannitol treatments and the only studied Ppc1hdz mutant presented a divergent function [45, 63]. These data suggest that the physiological role of C1HDZ genes in response to abiotic stress may be largely a vascular plant phenomenon.

MpC1HDZ-mediated depletion of oil bodies affects arthropod herbivory

We found compelling evidence that *MpC1HDZ* modulates, directly or indirectly, the differentiation of oil body cells. Consistently, loss-of-function *Mpc1hdz* alleles in two ecotypes exhibited a strong reduction in the number of oil bodies (Figure 1). Early experiments showed that the number and size of oil bodies in liverworts remained mostly unaffected in a variety of environmental conditions, such as light, temperature and limiting N or P [8, 64, 65], and that once formed they can be stable in herbarium specimens [10]. It was recently shown, however, that axenic growth resulted in a significant reduction in oil body cell formation, with a substantial induction upon a shift to non-axenic growth, such as cultivation on soil substrate where liverworts are challenged with biotic stressors [15]. We found that *Mpc1hdz* plants exhibit a severe reduction of oil body formation under all tested environmental conditions. However, the presence of a limited number of oil bodies in the mutants suggests that other transcriptional regulators could be relevant along with *MpC1HDZ*.

We explored several environmental conditions where oil bodies could play a physiological role. Whereas we found no evidence that *Mpc1hdz* plants were compromised in the response to an array of abiotic stressors, mutant plants were more susceptible to herbivory by the isopod *A. vulgare* (Figure 4C) and mutant plant extracts were less efficient controlling the growth of *B. subtilis* (Figure 4D). Thus, our results provide genetic evidence supporting a role for oil bodies in liverwort defenses against arthropod herbivores, supporting the hypothesis formulated more than a century ago [26].

What is the ancestral function of oil bodies?

Oil bodies are a synapomorphy of liverworts, and thus arose in the ancestral liverwort [7], likely placing their origin as early as the Ordovician and raising the question of their ancestral function. Whereas in liverworts where all cells contain oil bodies an antimicrobial role might be plausible, in liverworts with specialized oil body cells (idioblasts) a role in herbivory defense appears more likely. However, the ancestral condition in liverworts, scattered or ubiquitous, is equivocal. There is evidence from the fossil record that the selection pressure of herbivory was acting on ancestral liverworts. Middle Devonian liverwort fossils showing undigested tissues rich in opaque cells

resembling the oil bodies of extant liverworts provides indirect evidence of a role in deterring herbivory [27]. Whereas an earlier liverwort fossil record is missing, evidence of arthropod feeding on spores and stems of late Silurian-early Devonian plants has been reported [66]. Furthermore, evidence of herbivory corresponds closely with fossil evidence of plant organ evolution, implying either an immediate herbivory response or an extension of earlier plant-herbivore interactions. Thus, the anti-herbivory function of terpenoids and other secondary metabolites concentrated in oil bodies may date to the ancestral liverwort as was also recently suggested [67]. More broadly, as these compounds also play a central role in the ecology and evolution of the biotic response of seed plants [68], the origin and diversification of these secondary metabolites might have facilitated the evolution of early land plants via modulation of their interactions with the biotic environment.

Biochemical properties of *Mpc1hdz* mutants

As expected, the paucity of oil bodies in *Mpc1hdz* plants resulted in a marked reduction in specific terpene-related compounds (Figure 6). It has been demonstrated that the sesquiterpenoids and bis(bibenzyls) are accumulated in the oil bodies [15]. These compounds present cytotoxic activity against bacteria and fungi, mammalian and plant cells [14, 15, 29-31] but also have other medical applications [58]. Indeed, such bioactivity is the likely reason *M. polymorpha* is mentioned at least as early as a Greek herbal in the first century as a surface poultice to prevent inflammation of wounds [69]. *In planta*, while we provide evidence of a role for oil bodies against isopods, additional functions in fungal or bacterial resistance have not been ruled out. However, that liverwort species with oil bodies in every cell form mycorrhizal fungal interactions [70] indicates fungi can avoid their cytotoxicity, perhaps due to oil bodies being membrane bound organelles.

Oil body cell differentiation depends on MpC1HDZ

In *M. polymorpha* thalli, idioblasts emerge close to the apical notch in a proportion that later decreases due to unequal cell division rates [8]. Nonetheless, the precise mechanism of oil body differentiation

and formation is still unclear. Here, we described *Mpc1hdz* mutants that not only have a severe reduction of oil bodies but also do not differentiate idioblast cells (Figure 1). Transcriptional analysis of *Mpc1hdz* thalli demonstrated that several down-regulated genes encode proteins located in the oil body (TPS, MTPSL and MpSYP12B; [16, 60]), together with other secondary metabolism related genes that could be directly activated by MpC1HDZ (Figure 5). It is unclear whether the lethality encountered in *MpC1HDZ* gain-of-function mutants results from ectopic production of cytotoxic compounds, ectopic oil body cells, or another defect.

As oil bodies form at low frequency in *Mpc1hdz* mutants, other transcriptional regulators could be important for oil body cell differentiation. The R2R3-MYB TFs MpMYB02 and MpMYB14 control the bis(bibenzyl) and auronidin biosynthesis, respectively [35, 37]. *Mpc1hdz* mutants showed a strong repression of MpMYB02 expression, whereas MpMYB14, important for the defense against the oomycete pathogen *Phytophthora palmivora* [71], remained unchanged (Table S1, S2). That *MpC1HDZ* expression is not altered in MpMYB02 gain-of-function alleles [35] suggests that MpMYB02 is downstream MpC1HDZ activity, however, it is unknown whether MpMYB02 is also specifically expressed in oil body cells or participates in oil body cell differentiation and terpenoid biosynthesis. In contrast, we did not find any relationship of *MpC1HDZ* with MpMYCX/Y, which is associated with the defense-related induction of sesquiterpenoid biosynthesis in response to herbivory [67], or with ABA signaling, which is associated with bis(bibenzyl) increase under UV-C light stress [38].

Based on our observations, we propose a model for oil body cell differentiation (Figure 7). Undifferentiated cells derived from the notch express and accumulate MpC1HDZ to high levels; stage 1) a higher expression level of MpC1HDZ in one of the cells triggers differentiation of an oil body cell, while MpC1HDZ accumulation in neighboring cells decreases; stage 2) oil body forms in cells with high MpC1HDZ expression and oil body components are biosynthesized ; stage3) oil bodies mature while MpC1HDZ expression fades, and growth displaces them from the notch as non-idioblast

cells continue dividing (Figure 7). We propose that MpC1HDZ mediates the developmental program of oil body cell differentiation. Notably, our results do not exclude that other transcription factors may act earlier than MpC1HDZ in oil body cell differentiation. Investigating how integration of different transcriptional and environmental inputs promotes oil body cell development and differentiation will help with further understanding of this process. Such information will also pave the way for manipulating oil body numbers and composition for biotechnological purposes.

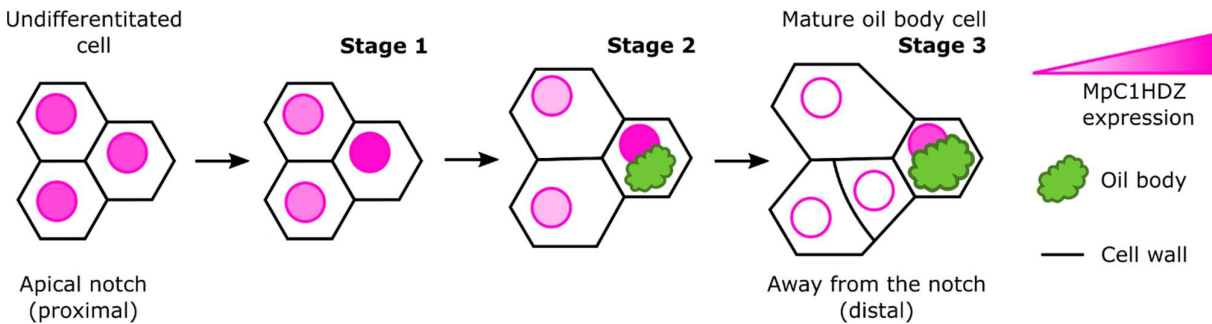


Figure 7. Proposed model for MpC1HDZ dynamics of oil body cell differentiation. All cells in proximity of an apical notch express MpC1HDZ. While most of the cells gradually lose MpC1HDZ expression, some cells start expressing MpC1HDZ stronger than their neighbors (stage 1). These cells can form an oil body (stage 2). With distance from apical notch and maturation of oil body cell, MpC1HDZ expression fades (stage 3). During oil body cell development, neighboring cells continue expanding and dividing, while oil body cells remain smaller.

MATERIAL AND METHODS

Plant material and growth conditions.

Marchantia polymorpha sp. *ruderalis* Australian (Melbourne) ecotype [72] and BoGa ecotype (obtained from Botanical Garden in Osnabrück, Germany) [73] were cultivated from gemmae under axenic conditions on vented 9 by 2.5 cm Petri dishes containing ½ strength Gamborg's B5 medium pH 5.5 solidified with 1 % (w/v) agar, referred as normal media. Australian ecotype was grown in growth chambers with 16/8 h photoperiod, white light at 80-100 $\mu\text{mol m}^{-2}\text{s}^{-1}$ at 22°C. BoGa ecotype

was grown in growth chambers with continuous white light at 50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ at 22°C. Sexual reproduction was induced using white light supplemented with far-red light (730 nm, Cree Inc. LEDs) as described before [74]. Plants were crossed in sterile conditions using autoclaved water for spermatozoid transfer. Crossing was repeated twice and plants were returned to long day condition until sporangia were ready for harvest. Sporangia were dried and stored at -80°C.

Marchantia transformation

M. polymorpha spore transformation was performed as described previously using both ecotypes [75]. Briefly, *M. polymorpha* spores were germinated and grown in 25 mL ½ strength liquid Gamborg's B5 with 2%(w/v) sucrose, 0.1%(w/v) Casamino acids and 0.03%(w/v) L-Glutamine for 10 days prior to co-cultivation with 1 mL of an *Agrobacterium tumefaciens* (GV3001) suspension, harboring binary plasmids, along with acetosyringone to a final concentration of 100 μM . Sporelings were plated on selection media (10 $\mu\text{g/mL}$ Hygromycin B with 200 mg/mL Timentin). After two weeks, T1 lines were transferred for one more week to new selection plates before being transferred to standard growth plates for gemmae production. Gemmae (G1 or G2 generation) from several independent primary transformants (T1 generation) were analyzed for the presence of the transgene by PCR and sequencing, with some selected for further analysis.

Translational reporters were transformed into *Mpc1hdz-3^{ge}* mutant using *Agrobacterium*-mediated transformation of regenerating thalli [76]. Briefly, plants were grown from gemmae on standard medium in continuous light for 14 days. 14 day-old thalli were cut into four pieces without apical notch and cultured for 3 days on standard medium with 1 % sucrose to induce regeneration. Regenerating thalli were co-cultured with *A. tumefaciens* GV3101 in liquid medium with 100 μM acetosyringone with agitation. After three days the plants were washed six times with autoclaved water, incubated for 30 min in 1 g/L cefotaxime and plated on selective plates containing 0.6 μM chlorsulfuron and 120 $\mu\text{g/mL}$ cefotaxime. After transformants appeared, they were transferred to new selective plates. G2 generation was analyzed for fluorescence and presence of the transgene by PCR.

Genetic constructs

All primer sequences are available in Table S5.

gRNA design and cloning

Synthetic guide RNAs (gRNA) targeting genes of interest were co-expressed on independent T-DNAs using pGE010 plasmid [77, 78]. Design of specific gRNAs was performed such that putative off-targets (*M. polymorpha* genome assembly v3.1) lacked an associated protospacer adjacent motif (PAM) site. For single gRNAs transformation, primers MpC1HDZ_gRNA1 were cloned downstream of the MpU6 promoter, excluding PAM sites, in pMpGE_En03 using BsaI sites and were co-expressed with Cas9 [77]. For double gRNAs transformations, gRNAs were designed using the online tool <http://marchantia.info/casfinder>. The first gRNAs (gRNA-wg-1 or gRNA-HD-1) were cloned into pMpGE_En04 using BsaI sites [79]. Second gRNAs (gRNA-wg-2 or gRNA-HD-2) were cloned into pMpBC_GE14 using BsaI sites. Both vectors were digested using BglI. The pMpGE_En04 backbone with the first gRNA and the fragment with the second gRNA were ligated, resulting in pMpGE_En04 with both gRNAs in tandem each under the control of the MpU6 promoter. These vectors were used in a Gateway LR reaction with the binary vector pMpGE010 featuring Cas9 under the control of the EF1 promoter [77-79].

MpC1HDZ cloning and genetic constructs

The MpC1HDZ (*Mp3g02320*) coding sequence from *M. polymorpha* Tak-1 genome v3.1 (marchantia.info) was used to design specific primers. The cDNA was subcloned into pENTR-DTOPO (Thermo Fisher, Waltham, MA, US; pENTR-MpC1HDZ) using primers MpC1HDZ_CACC_fw and MpC1HDZ_rv. The CDS was mutagenized using primers MpC1HDZ_gibson_fw and MpC1HDZ_mut_rv, and MpC1HDZ_mut_fw and MpC1HDZ_gibson_rv in order to make it resistant to the gRNA. Fragments were purified and used as template for a subsequent amplification using primers MpC1HDZ_gibson_fw and

509 MpC1HDZ_gibson_rv and cloned in a pENTR2B (Thermo Fisher, Waltham, MA, US) using
510 NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs Inc., UK, pE2B-MpC1HDZ_{mut}).
511 Later, it was subsequently recombined using the Gateway LR Clonase II Enzyme Mix (Thermo
512 Fisher, Waltham, MA, US) according to the manufacturer's instructions into different destination
513 vectors: pMpGWB403 (*proEF1:MpC1HDZ_{mut}*), pMpGWB406 (*pro35S:MpC1HDZ_{mut}-Citrine*) [80].
514 As a control, a 4.6 kb promoter region of MpC1HDZ was amplified with primers
515 *proMpC1HDZ_gibson_fw* and *proMpC1HDZ_gibson_rv* as described below. Both fragments were
516 used for cloning in a pENTR2B with NEBuilder HiFi DNA Assembly Master Mix and recombined
517 in pGWB401 (*proMpC1HDZ^{4.6kb}-MpC1HDZ_{mut}*).
518 For overexpression in tobacco, pENTR-MpC1HDZ was recombined in the pFK247 expression vector
519 derived from the pGreen vector series [81] for transient expression in tobacco plants.

520 **Transcriptional reporters**

521 The promoter sequence of MpC1HDZ was obtained from Phytozome. A 4.6 kb fragment of the
522 promoter was amplified using primers *proMpC1HDZ-NdeI-fw* and *proMpC1HDZ-SalI-rv* and cloned
523 into pCRII-TOPO. It was subsequently subcloned into pRITA, which has a multiple cloning site
524 upstream of a GUS/UidA reporter gene and *nos* terminator [82], using NdeI and KpnI. Finally,
525 *proMpC1HDZ^{4.6kb}:GUS* was cloned into pSKF-KART using NotI [72].

526 Alternatively, a 3 kb sequence upstream of MpC1HDZ start was amplified from BoGa DNA using
527 primers with attB1 and attB2 sites. The subsequent PCR product was recombined with pDONR201
528 in BP reaction. The entry clone was recombined in an LR reaction with pMpGWB104 to obtain GUS
529 reporter [80].

530 **Translational reporters**

531 To obtain translational reporters, cDNA was obtained from total RNA extracted from BoGa ecotype
532 thalli. MpC1HDZ was amplified from cDNA using a specific forward primer and polyT reverse

primer (Roche) and ligated into pGEMT. The amplified CDS starts at position 3 and ends at position 720 relative to the genome prediction and therefore likely represents a shorter transcript. Mp*C1HDZ* CDS was amplified again from this plasmid to add *SpeI* sites. The 3 kb sequence upstream of Mp*C1HDZ* CDS start was amplified from BoGa DNA and ligated into pGEMT vector. Mp*C1HDZ* CDS was inserted behind the 3 kb Mp*C1HDZ* promoter using *SpeI* restriction sites. The 3 kb Mp*C1HDZ* promoter and Mp*C1HDZ* CDS combination were amplified using primers with attB1 and attB2 sites and recombined in BP reaction with pDONR201. The entry clone was recombined in LR reaction with pMpGWB307 destination vector to obtain citrine fusion, or with pMpGWB329 to obtain tdTomato fusion [80].

Histology and microscopy

Fresh *M. polymorpha* gemmae were resuspended in 0.1% (v/v) Triton-X100 and observed using a compound microscope (Nikon Eclipse E200, Nikon Instruments Inc., Melville, NY, USA; or Zeiss Axio Imager M2) or Leica MZ10F stereomicroscope (Leica Microsystems, Wetzlar, Germany; or Nikon SMZ1500). In the case of thallus and cup cross sections, fresh thalli were sliced into 0.5-0.8 mm thick slices manually using a razor blade and observed in the compound microscope.

Oil body staining was performed using a 10 mg mL⁻¹ Nile Red dissolved in phosphate buffered saline (PBS) just before use as described before [15]. This solution was vacuum infiltrated for 1 min, incubated 10 min at room temperature and then rinsed with 0.1% (w/v) Triton-X100 solution. For BODIPY staining, the stock solution was diluted 1000 times in PBS to prepare the working solution. Plants were covered with BODIPY for 10 min in darkness and washed three times with PBS.

The samples were observed using GFP filter under fluorescent microscope (Zeiss Axio Imager M2). Oil bodies in gemmae were manually counted from images. Thallus area was quantified at different time intervals from pictures with ImageJ software. Thallus growth rate is the proportion between final and initial area of the thallus. All data was analyzed and plotted in R with built-in packages.

GUS staining

GUS assays were performed as described previously [72]. Plants were immersed in GUS staining solution (100 mM Potassium Ferrocyanide, 100 mM Potassium Ferricyanide and 1 mM X-Gluc) and after applying 2 rounds of vacuum for 1 min, the plants were incubated for 1.5 h to 2 h at 37°C in darkness, cleared with ethanol, and imaged using a compound microscope or stereoscopic microscope. BoGa ecotypes were stained similarly but with slightly different GUS staining solution composition (50 mM phosphate buffer, 10 mM EDTA, 50 mM Potassium Ferrocyanide, 50 mM Potassium Ferricyanide and 1.4 mM X-Gluc) and the stain was vacuum-infiltrated for 10 min before incubation at 37 °C.

Oil body cell and oil body distance measurement

For live imaging, gemmalings were grown on standard medium in small Petri dishes (6 cm diameter). 3 day-old gemmalings were stained with BODIPY 493/503 working solution (dissolved in water) for 3 min. Approximately 30 µL dye was used to cover a gemmaling. Samples were rinsed with water two times, covered with water and observed using Leica TCS SP8 confocal microscope and 25x 0.95 NA water dip-in objective. To observe fluorescence, 561 nm excitation and 570 - 600 nm emission was used for tdTomato and 488 nm excitation and 503 – 533 nm emission was used for BODIPY in sequential scan mode. All samples were imaged using the same settings. Images were analyzed in Fiji software. First, all stacks were trimmed to the same size relative to the estimated position of the apical cell along the z-axis. Prior to applying “sum slices” projection to the trimmed stacks, black and white pixels were added in a lower left corner of each slice as a calibration for lowest and highest possible values. In these projections, some nuclei visually stood out as much brighter than the surrounding nuclei. To visualize only these nuclei, all projections were processed as follows: format was set to 16 bit, intensities were multiplied by 100 and smoothed once. For tdTomato channel, brightness & contrast values were set to show only the pixels with intensity value of 10 000 or more. To assess the distances of bright nuclei and oil bodies from the apical notch, distance bins were

created using Concentric circles plug-in. Circles were centered in the middle of the apical notch of each image and placed in 25 μm distance intervals. Oil bodies and bright nuclei were quantified for each of the distance bins. Data was visualized and statistically analyzed in R. To visualize apical notch surface in 3D, a surface mesh was created in MorphoGraphX program using signal from BODIPY channel [83].

Oil body cell tracking

For live imaging, gemmalings were grown on normal media in small Petri dishes (6 cm diameter). A few droplets of water were added around the rim of the medium 30 min before imaging, to soak the agar and prevent sample movements during the imaging. Gemmalings were stained with BODIPY 493/503 working solution (dissolved in water) for 5 min. Approximately 30 μL dye was used to cover a 5-day-old gemmaling. 15 μL of 0.1 % propidium iodide (PI) was added directly to BODIPY and removed after 30 s. Samples were rinsed with water three times, covered with water and observed using Leica TCS SP8 confocal microscope and 25x 0.95 NA water dip-in objective. To observe fluorescence, 488 nm excitation and 503 – 523 nm emission was used for BODIPY, 552 nm excitation and 560 – 580 nm emission was used for tdTomato, and 552 nm excitation and 620 – 640 nm emission was used for PI. The two lasers were used in sequential scan (between stacks), and Lightning mode was used for acquisition. Standard deviation projections of z-stacks were generated in Fiji and used for oil body cell tracking. In case of doubt about the identity or lineage of an object, the stacks were observed in MorphoGraphX.

Plant stress treatments

Wild-type and *Mpc1hdz* mutant gemmae were grown in agar plates containing normal media. After 1-week growth in control conditions, each plant was transferred with tweezers to a new plate or pot to be exposed to different treatments. Plant areas were recorded before and after the treatment to estimate the impact on the relative growth rate of each genotype. Plant growth rate is the proportion between final and initial area of the thallus. Pictures of whole plants were taken at different time

intervals using a Leica MZ10F stereomicroscope. Thallus area was quantified from pictures with ImageJ software. For cold treatments and high light intensity (HLI) plants were transferred to plates containing the normal media (3 plates with 12 plants for treatment and 3 plates as control). Control plates remained in control conditions. Cold treatment was performed at 4°C in a chamber with similar light and photoperiod conditions and observed after 1 week. HLI treatment was performed in the same chamber but 1000 $\mu\text{mol m}^{-2} \text{s}^{-1}$ LED lights with the same photoperiod and observed after 1 week. For UV treatment plants were transferred to normal media and exposed to a laminar flow bench UV-C light for 1 h with the lid open (UV+) or closed (UV-), after that they were recovered in control conditions and observed after 1-week. For osmotic stress, normal media was supplemented with solutions of NaCl or mannitol at a 50mM final concentration or water (mock). Plants were grown in control conditions for 18 days. For starvation, plants were transferred to 1/100 Gamborg's B5 media (1% (w/v) agar, pH 5.5), grown in control conditions, and observed after 2 weeks (normal media was used as control). For non-axenic growth, plants were transplanted to pots with rockwool as substrate in a growth chamber in similar temperature and photoperiod as control conditions with irrigation. The water loss experiment was performed as in Re *et al.* [84]. Briefly, 15-day-old wild-type and *Mpc1hdz* plants grown in control conditions and normal media were detached from the agar, placed in a paper towel, and weighed every 30 min during 2 h at room temperature. The percentage (%) of water loss was calculated as follows: $[\text{initial weight} - \text{weight}] / [\text{initial weight}] \times 100$.

Pill bug feeding assay

Feeding assay was performed according to Nakazaki *et al.*, with slight modification [85]. Gemmae were cultured on 1/2 Gamborg's B5 medium (Nihon Pharmaceutical, Tokyo, Japan, Cat#399-00621) containing 1%(w/v) agar (Nacalai tesque, Kyoto, Japan, Cat#01028-85) and 1%(w/v) sucrose (Wako Pure Chemical, Osaka, Japan, Cat#196-00015) for 5 days at 22°C under continuous white fluorescent light (50 $\mu\text{mol m}^{-2}\text{s}^{-1}$). Five-day-old thalli were transferred onto 1/2× Gamborg's B5 medium containing 1%(w/v) agar without sucrose and cultured for additional 5 days under the same condition

above, or 8 days in case of BoGa ecotype. Pill bugs (*Armadillidium vulgare*) were collected at the Myodaiji area of National Institute for Basic Biology (Aichi, Japan). Before the assay, pill bugs were maintained on Prowipe (Daio Paper, Tokyo, Japan) moistened with sterilized water for 48 h at 22°C under dark condition to be starved. Six pill bugs were introduced into each medium plate containing the 10-day-old thalli, or 13-day-old thalli in case of BoGa ecotype, and kept together for 24 h under the dark condition at 22°C. The thallus areas of *M. polymorpha* were calculated using ImageJ (National Institute of Health, <https://imagej.nih.gov/ij/>), and figures were processed with Photoshop software (Adobe systems, USA).

Ethanol extraction and overlay assay

Wild-type and *Mpc1hdz* gemmae were grown in agar plates containing normal media and control then transferred to pots using soil and vermiculite (1:1 parts) and grown in control conditions with irrigation for one month. Plant materials were snap-frozen with liquid N₂, and ground using a mortar and pestle (~5 g fresh weight). The plant powder was extracted with EtOH for 24 h in agitation. Debris was removed by filtration and removal of solvent was performed using a SpeedVac. The stock solution of crude extracts was prepared as 10 mg mL⁻¹ in DMSO. *Bacillus subtilis* overlay assay was performed as described before [86]. For quantification purposes, pictures of bacterial growth inhibition were taken using stereoscopic microscope (Leica MZ10F) and calculated as density of colonies.

RNA-seq

Total RNA isolated from 3-week-old thalli (QIAGEN RNeasy Plant Kit 5) from each genotype was assessed for quality with a Nanodrop spectrophotometer and a Bioanalyzer 2100 microfluidics system (Agilent 6). For expression analysis in control conditions, one sample per genotype was used (*Mpc1hdz-1^{ge}*, *Mpc1hdz-2^{ge}*, WT female, WT male) and mutant and wild-type were treated as replicates. For expression in non-axenic conditions, plants were grown as described before and three biological replicates per genotype were taken from *Mpc1hdz-2^{ge}* and WT female. Library preparation

657 substrates were enriched for mRNA (Illumina TruSeq Stranded mRNA technology). Sequencing used
658 the Illumina NextSeq500 7 in HighOutput mode (20 million single-ended 75 b reads per sample).
659 Reads were mapped onto the Marchantia v3.1 (<http://marchantia.info/>) assembly using TopHat v2.1.0
660 for Galaxy [87]. Resulting BAM files were loaded onto IGV 1.3.1 to produce Sashimi Plots [88]. A
661 count matrix of average raw reads/gene/sample was created using HTseq-code using Galaxy [89].
662 Raw reads were normalized by total million reads per library (RPM) and transcript lengths in kb
663 (RPKM). The edgeR package was used for differential gene expression analysis using Galaxy (Table
664 S1 and S2) [90]. Summary statistics of RNA-seq analysis are available in Table S6.

665 For GO term and protein family enrichment analysis, the algorithm and annotations were as described
666 previously [91]. Gene names from Marchantia v5.1 [92] was assigned in accordance to gene
667 corresponding table available at <http://marchantia.info/nomenclature/> [2]. Genes were annotated
668 using information nomenclature previously published [93] or available at MarpolBase.

669 **Hexane extraction and gas chromatography**

670 Samples of WT and transgenic plants (3-4 samples per genotype) were collected from 3-week-old
671 plants in different treatments and snap frozen in liquid nitrogen. Each sample was ground with a
672 pestle in a mortar containing liquid nitrogen. 300 μ L of n-hexane was added to ~400 mg of ground
673 tissue, centrifuged and extracted overnight with shaking at 100 rpm. For samples in starvation,
674 extraction was scaled-down to 50 μ L of solvent and ~50 mg of grounded tissue. Each extract was
675 then vortexed and centrifuged, and the supernatant was transferred into glass vials for gas
676 chromatographic analyses.

677 Constituents in hexane extracts were identified by GC-MS and quantified by GC-FID. The GC-MS
678 system comprised a Gerstel 2.5.2 autosampler, a 7890A gas chromatograph and a 5975C quadrupole
679 MS (Agilent Technologies, Santa Clara, USA). Aliquots of extracts (1 μ L) were injected in splitless
680 mode and the MS was adjusted to manufacturer's recommendations using tris-(perfluorobutyl)-amine
681 (FC-43). We used the following MS source conditions: injection temperature 250°C, transfer line

280°C, ion source 230°C, quadrupole 150°C, 70 eV (EI mode), 2.66 scans s⁻¹ and scanning range *m/z* 50-600. The GC-MS column was a VF-5MS (30 m × 250 µm i.d. with 0.25 µm film thickness) fitted with a 10 m EZ-Guard column (J & W, Agilent). Helium was used as the carrier gas at a flow rate of 1 mL min⁻¹. The column temperature was held at 70°C for 6 min following injection, then ramped at 10°C min⁻¹ to 320°C and held at that temperature for 5 min. Mass spectra were evaluated using Agilent MSD ChemStation (E.02.02.1431) and sesquiterpenes were identified using a mass spectra library described before [94]. Other constituents were identified using the NIST 11 database (<https://chemdata.nist.gov/dokuwiki/doku.php?id=chemdata:start>). Constituent identity was confirmed by their relative retention times and by comparison to authentic standards where available, including the FAME mix (Sigma-Aldrich, St. Louis, USA).

Quantification of constituents in extracts was performed using a GC-FID system comprising a Perkin Elmer Autosystem XC (Perkin Elmer, Melbourne, Australia) fitted with a low polarity Zebron ZB-5 (30 m × 250 µm i.d., Phenomenex) column. Aliquots of extracts (2 µL) were injected under a split flow of 7.5 mL He min⁻¹. The injector and detector temperatures were 275 and 320°C, respectively. The temperature program was the same as for the GC-MS analyses. Quantification of constituents in each class was based on calibration series of commercial standards of d-limonene (monoterpenes and C₈ volatiles), aromadendrene (aromadendrane sesquiterpene hydrocarbons), valencene (other bicyclic sesquiterpene hydrocarbons), caryophyllene oxide (oxygenated sesquiterpenes), hexadecane (hydrocarbons), palmitic acid (fatty acids and aldehydes), phytol (diterpenes) and stigmasterol (phytosterols; Sigma-Aldrich).

Statistical analysis

Statistical significance was calculated using ANOVA from agricolae package using Rstudio software (<https://www.rstudio.com>) and corrected by Tukey HSD (alpha = 0.05) for levels calculations. In other cases, built-in packages in R were used for pairwise *t*-test, Wilcoxon rank sum test and binominal test. In each figure it is explained which test was used.

ACKNOWLEDGEMENTS

We thank Dr. Teruyuki Niimi (National Institute for Basic Biology) and Dr. Takahisa Miyatake (Okayama University) for identification of *Armadillidium vulgare*. The support of plant cultivation rooms was provided by the Model Plant Research Facility of National Institute for Basic Biology. We thank Dr. Magnus Eklund (Uppsala University) for his help with transcriptional reporters. We thank Dr. Detlef Weigel (Max Planck Institute, Germany) for kindly providing pFK247 expression vector, Dr. Kimitsune Ishizaki (Kobe University) for the cloning vectors and Dr. Keisuke Inoue for providing pMpGE_En04 and pBC-14. We thank Dr. Saiko Yoshida and Dr. Peter Huijser for assistance with confocal imaging and image analysis and Michelle van der Gragt for GUS stainings of BoGa plants.

FUNDING

This work was supported by Agencia Nacional de Promoción Científica y Tecnológica (PICT2017-1484 to J.E.M.) and Conicet (PIP267 to J.E.M.). S.N.F. and J.L.B. were supported by the Australian Research Council (DP170100049). This work was financially supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (to T.U., 17K19412, 18H02470 and 19H05675, and T.K., 17H07333 and 18K14738), The Mitsubishi Foundation, and Yamada Science Foundation. This work was supported by a core grant from the Max Planck Society (M.T.). SZ acknowledges support from the Deutsche Forschungsgemeinschaft (SFB 944). F.R. is a doctoral fellow of CONICET. J.E.M. is a CONICET career member.

AUTHOR CONTRIBUTIONS: FR, EB, JLB, MT and JEM designed the research; MT and JEM conceived the project; FR, EB, SNF, JQD, RM, TD, SZ, TK and TU performed research, FR, EB, JLB, MT and JEM wrote the paper.

SUPPLEMENTARY INFORMATION

Figure S1. MpC1HDZ resembles structural features of angiosperms C1HDZ orthologs.

Figure S2. Plant morphology of wild-type and mutant alleles.

Figure S3. Plant growth responses upon abiotic stress.

Figure S4. Constitutive overexpression of MpC1HDZ may be lethal.

Figure S5. MpC1HDZ promoter activity in BoGa ecotype.

Figure S6. MpC1HDZ promoter activity upon abiotic stress treatments.

Figure S7. Measurement of distance from apical notch to nuclei highly expressing MpC1HDZ and oil bodies.

Figure S8. Overview of oil body cell tracking time lapse.

Figure S9. Thalli growth rate of *Mpc1hdz* under different abiotic stress treatments.

Figure S10. Expression of *HDZ* genes under ABA and abiotic stress treatments.

Figure S11. Pill bug feeding assay on mutant alleles of the BoGa ecotype.

Figure S12. Dose-dependence effect of plant extract on the growth of *Bacillus subtilis*.

Figure S13. Representative figures of plants after treatments.

Figure S14. Mass spectrum of unidentified compound U13.

Table S1. Differentially expressed genes upon axenic conditions in *Mpc1hdz*.

Table S2. Differentially expressed genes upon non-axenic conditions in *Mpc1hdz*.

752 Table S3. GO categories for down-regulated genes in *Mpc1hdz*.

753 Table S4. Chemical compounds identification in hexane extracts.

754 Table S5. Primer list.

755 Table S6. Summary statistics of RNA-seq experiments.

756

757 LITERATURE

- 758 1. Delwiche, C.F., and Cooper, E.D. (2015). The evolutionary origin of a terrestrial flora. *Curr.*
759 *Biol.* 25, R899-910.
- 760 2. Bowman, J.L., Kohchi, T., Yamato, K.T., Jenkins, J., Shu, S., Ishizaki, K., Yamaoka, S.,
761 Nishihama, R., Nakamura, Y., Berger, F., et al. (2017). Insights into land plant evolution
762 garnered from the *Marchantia polymorpha* genome. *Cell* 171, 287-304.
- 763 3. Catarino, B., Hetherington, A.J., Emms, D.M., Kelly, S., and Dolan, L. (2016). The stepwise
764 increase in the number of transcription factor families in the Precambrian predated the
765 diversification of plants on land. *Mol. Biol. Evol.* 33, 2815-2819.
- 766 4. de Mendoza, A., Sebe-Pedros, A., Sestak, M.S., Matejcic, M., Torruella, G., Domazet-Loso,
767 T., and Ruiz-Trillo, I. (2013). Transcription factor evolution in eukaryotes and the assembly of
768 the regulatory toolkit in multicellular lineages. *Proc. Natl. Acad. Sci. USA* 110, E4858-4866.
- 769 5. Morris, J.L., Puttick, M.N., Clark, J.W., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H.,
770 Yang, Z., Schneider, H., and Donoghue, P.C.J. (2018). The timescale of early land plant
771 evolution. *Proc. Natl. Acad. Sci. USA* 115, E2274-E2283.
- 772 6. Puttick, M.N., Morris, J.L., Williams, T.A., Cox, C.J., Edwards, D., Kenrick, P., Pressel, S.,
773 Wellman, C.H., Schneider, H., Pisani, D., et al. (2018). The interrelationships of land plants
774 and the nature of the ancestral Embryophyte. *Curr. Biol.* 28, 733-745 e732.
- 775 7. Schuster, R.M. (1966). *The Hepaticae and Anthocerotae of North America, Volume I*, (New
776 York: Columbia University Press).
- 777 8. Slavík, B. (1950). Specific bodies in the cells of hepatics (Hepaticae). *Stud. Bot. Cech* 11, 226-
778 244.
- 779 9. Gottsche, C.M. (1843). Anatomisch-physiologische untersuchungen über *Haplomitrium*
780 *hookeri* N.v.E., mit vergleichung anderer lebermoose. *Nov. Actorum Acad. Caes. Leop.-Carol.*
781 *Nat. Cur.* 20: 267–398.
- 782 10. Lindberg, S.O. (1862). Ueber das vorkommen von aetherischen oelen in lebermoosen. *Flora*
783 45, 545-546.
- 784 11. Lohmann, C.E.J. (1903). Beitrag zur chemie und biologie der leber moose. *Botan. Centrallbl.*
785 *Beih.* 15, 215–256.
- 786 12. Müller, K. (1905). Beitrag zur kenntnis der ätherischen öle bei lebermoosen. In Hoppe-Seyler's
787 *Zeitschrift für physiologische Chemie*, 45: 299.
- 788 13. Asakawa, Y., and Ludwiczuk, A. (2018). Chemical constituents of bryophytes: structures and
789 biological activity. *J. Nat. Prod.* 81, 641-660.

- 790 14. Chen, F., Ludwiczuk, A., Wei, G., Crandall-Stotler, B., and Bowman, J.L. (2018). Terpenoid
791 secondary metabolites in bryophytes: chemical diversity, biosynthesis and biological functions.
792 Crit. Rev. Plant Sci. 37.
- 793 15. Tanaka, M., Esaki, T., Kenmoku, H., Koeduka, T., Kiyoyama, Y., Masujima, T., Asakawa, Y.,
794 and Matsui, K. (2016). Direct evidence of specific localization of sesquiterpenes and
795 marchantin A in oil body cells of *Marchantia polymorpha* L. Phytochemistry 130, 77-84.
- 796 16. Suire, C., Bouvier, F., Backhaus, R.A., Begu, D., Bonneu, M., and Camara, B. (2000). Cellular
797 localization of isoprenoid biosynthetic enzymes in *Marchantia polymorpha*. Uncovering a new
798 role of oil bodies. Plant Physiol. 124, 971-978.
- 799 17. Pfeffer, W. (1874). Die oelkörper der lebermoose. Flora 57, 2-6; 17-27; 33-43; Tafel I.
- 800 18. Pihakaski, K. (1968). A study of the ultrastructure of the shoot apex and leaf cells in two
801 liverworts, with special reference to the oil bodies. Protoplasma 66, 79-103.
- 802 19. Suire, C. (2000). A comparative, transmission-electron microscopic study on the formation of
803 oil-bodies in liverworts. J. Hattori Bot. Lab. 89, 209-232.
- 804 20. Dombrey, P. (1926). Contribution a l'étude des corps oléiformes des hépatiques des environs
805 de Nancy. (Paris), p. 211.
- 806 21. Gavaudan, P. (1930/1). Recherches sur la cellule des hépatiques. Le Botaniste 22, 105-294.
- 807 22. Pressel, S., Duckett, J.G., Ligrone, R., and Proctor, M.C.F. (2009). Effects of de- and
808 rehydration in desiccation-tolerant liverworts: a cytological and physiological study. Int. J.
809 Plant Sci. 170, 182-199.
- 810 23. Hieronymus, G. (1892). Beiträge zur morphologie und biologie der algen. Beiträge zur
811 Biologie der Pflanzen 5, 461-495; Taf. XVII-XVIII.
- 812 24. Chalaud, G. (1931). Germination des spores et formation du game'tophyte chez *Lophocolea*
813 *cuspidata* et *Chiloscyphus polyanthus*. Ann Bryol 4, 49-78.
- 814 25. Gavaudan, P. (1927). Sur l'origine et les caractères des éléments oléifères des Jungermaniales.
815 Comptes rendus hebdomadaires des séances de l'Académie des sciences 184, 1473-1475.
- 816 26. Stahl, E. (1888). Pflanzen und schnecken: biologische studie über die schutzmittel der pflanzen
817 gegen schneckenfrass. Jena Z. Med. Naturwiss. 22, 557-684.
- 818 27. Labandeira, C.C., Trembay, S.L., Bartowski, K.E., and Hernick, L.V. (2014). Middle Devonian
819 liverwort herbivory and antiherbivore defence. New Phytol. 202, 247-258.
- 820 28. Gahtori, D., and Chaturvedi, P. (2011). Antifungal and antibacterial potential of methanol and
821 chloroform extracts of *Marchantia polymorpha* L. Arch. Phytopathol. Plant Protect. 44, 726-
822 731.
- 823 29. Asakawa, Y. (1990). Biologically active substances from bryophytes. In Bryophyte
824 Development: Physiology and Biochemistry, R.N. Chopra, Bhatla, S.C., ed. (Boca Raton: CRC
825 Press), pp. 259-287.
- 826 30. Banerjee, R.D., and Sen, S.P. (1979). Antibiotic activity of bryophytes. Bryologist 82, 141-
827 153.
- 828 31. Komala, I., Ito, T., Yagi, Y., Nagashima, F., and Asakawa, Y. (2010). Volatile components of
829 selected liverworts, and cytotoxic, radical scavenging and antimicrobial activities of their crude
830 extracts. Nat. Prod. Commun. 5, 1375-1380.
- 831 32. Wu, Y.F., Zhao, Y., Liu, X.Y., Gao, S., Cheng, A.X., and Lou, H.X. (2018). A bHLH
832 transcription factor regulates bisbibenzyl biosynthesis in the liverwort *Plagiochasma*
833 *appendiculatum*. Plant Cell Physiol. 59, 1187-1199.

33. Kumar, S., Kempinski, C., Zhuang, X., Norris, A., Mafu, S., Zi, J., Bell, S.A., Nybo, S.E., Kinison, S.E., Jiang, Z., et al. (2016). Molecular diversity of terpene synthases in the liverwort *Marchantia polymorpha*. *Plant Cell* 28, 2632-2650.
34. Peters, K., Gorzolka, K., Bruelheide, H., and Neumann, S. (2018). Seasonal variation of secondary metabolites in nine different bryophytes. *Ecol. Evol.* 8, 9105-9117.
35. Kubo, H., Nozawa, S., Hiwatashi, T., Kondou, Y., Nakabayashi, R., Mori, T., Saiti, K., Takanashi, K., Kohchi, T., and Ishizaki, K. (2018). Biosynthesis of riccionidins and marchantins is regulated by R2R3- MYB transcription factors in *Marchantia polymorpha*. *J. Plant Res.* 131, 849-864.
36. Arai, H., Yanagiura, K., Toyama, Y., and Morohashi, K. (2019). Genome-wide analysis of MpBHLH12, a IIIIf basic helix-loop-helix transcription factor of *Marchantia polymorpha*. *J. Plant Res.* 132, 197-209.
37. Albert, N.W., Thrimawithana, A.H., McGhie, T.K., Clayton, W.A., Deroles, S.C., Schwinn, K.E., Bowman, J.L., Jordan, B.R., and Davies, K.M. (2018). Genetic analysis of the liverwort *Marchantia polymorpha* reveals that R2R3MYB activation of flavonoid production in response to abiotic stress is an ancient character in land plants. *New Phytol* 218, 554-566.
38. Kageyama, A., Ishizaki, K., Kohchi, T., Matsuura, H., and Takahashi, K. (2015). Absciscic acid induces biosynthesis of bisbibenzyls and tolerance to UV-C in the liverwort *Marchantia polymorpha*. *Phytochemistry* 117, 547-553.
39. Romani, F., Reinheimer, R., Florent, S.N., L., B.J., and Moreno, J.E. (2018). Evolutionary history of HOMEODOMAIN LEUCINE ZIPPER transcription factors during plant transition to land. *New Phytol.* 219, 408-421.
40. Romani, F., Ribone, P.A., Capella, M., Miguel, V.N., and Chan, R.L. (2016). A matter of quantity: common features in the drought response of transgenic plants overexpressing HD-Zip I transcription factors. *Plant Sci.* 251, 139-154.
41. Cabello, J.V., Arce, A.L., and Chan, R.L. (2012). The homologous HD-Zip I transcription factors HaHB1 and AtHB13 confer cold tolerance via the induction of pathogenesis-related and glucanase proteins. *Plant J.* 69, 141-153.
42. Gonzalez-Grandio, E., Pajoro, A., Franco-Zorrilla, J.M., Tarancon, C., Immink, R.G., and Cubas, P. (2017). Absciscic acid signaling is controlled by a BRANCHED1/HD-ZIP I cascade in *Arabidopsis* axillary buds. *Proc. Natl. Acad. Sci. USA* 114, E245-E254.
43. Ribone, P.A., Capella, M., and Chan, R.L. (2015). Functional characterization of the homeodomain leucine zipper I transcription factor AtHB13 reveals a crucial role in *Arabidopsis* development. *J. Exp. Bot.* 66, 5929-5943.
44. Moreno, J.E., Romani, F., and Chan, R.L. (2018). *Arabidopsis thaliana* homeodomain-leucine zipper type I transcription factors contribute to control leaf venation patterning. *Plant Signal. Behav.* 13, e1448334.
45. Sakakibara, K., Nishiyama, T., Sumikawa, N., Kofuji, R., Murata, T., and Hasebe, M. (2003). Involvement of auxin and a homeodomain-leucine zipper I gene in rhizoid development of the moss *Physcomitrella patens*. *Development* 130, 4835-4846.
46. Bowman, J.L., Araki, T., Arteaga-Vazquez, M.A., Berger, F., Dolan, L., Haseloff, J., Ishizaki, K., Kyoizuka, J., Lin, S.S., Nagasaki, H., et al. (2016). The naming of names: guidelines for gene nomenclature in *Marchantia*. *Plant Cell Physiol.* 57, 257-261.
47. Zhao, Y., Ma, Q., Jin, X., Peng, X., Liu, J., Deng, L., Yan, H., Sheng, L., Jiang, H., and Cheng, B. (2014). A novel maize Homeodomain-Leucine Zipper (HD-Zip) I gene, *Zmhdz10*, positively regulates drought and salt tolerance in both rice and *Arabidopsis*. *Plant Cell Physiol.* 55, 1142-1156.

- 881 48. Gago, G.M., Almoguera, C., Jordano, J., Gonzalez, D.H., and Chan, R.L. (2002). *Hahb-4*, a
882 homeobox-leucine zipper gene potentially involved in abscisic acid-dependent responses to
883 water stress in sunflower. *Plant Cell Environ.* 25, 633-640.
- 884 49. Cabello, J.V., Giacomelli, J.I., Gomez, M.C., and Chan, R.L. (2017). The sunflower
885 transcription factor HaHB11 confers tolerance to water deficit and salinity to transgenic
886 *Arabidopsis* and alfalfa plants. *J. Biotechnol.* 257, 35-46.
- 887 50. Ferreira, M.A., Pinheiro Da Cruz Waltenberg, F., Romano De Campos Pinto, E., and Grossi de
888 Sa, M.F. (2013). Use of the coffee homeobox gene *CAHB12* to produce transgenic plants with
889 greater tolerance to water scarcity and salt stress. (United States: Universidade Federal do Rio
890 de Janeiro, Rio de Janeiro (BR)).
- 891 51. Bhattacharjee, A., Khurana, J.P., and Jain, M. (2016). Characterization of rice homeobox genes,
892 *OsHOX22* and *OsHOX24*, and over-expression of *OsHOX24* in transgenic *Arabidopsis* suggest
893 their role in abiotic stress response. *Front. Plant Sci.* 7, 627.
- 894 52. Olsson, A.S., Engstrom, P., and Soderman, E. (2004). The homeobox genes *ATHB12* and
895 *ATHB7* encode potential regulators of growth in response to water deficit in *Arabidopsis*. *Plant*
896 *Mol. Biol.* 55, 663-677.
- 897 53. Valdes, A.E., Overnas, E., Johansson, H., Rada-Iglesias, A., and Engstrom, P. (2012). The
898 homeodomain-leucine zipper (HD-Zip) class I transcription factors *ATHB7* and *ATHB12*
899 modulate abscisic acid signalling by regulating protein phosphatase 2C and abscisic acid
900 receptor gene activities. *Plant Mol. Biol.* 80, 405-418.
- 901 54. Ghosh, T.K., Kaneko, M., Akter, K., Murai, S., Komatsu, K., Ishizaki, K., Yamato, K.T.,
902 Kohchi, T., and Takezawa, D. (2016). Abscisic acid-induced gene expression in the liverwort
903 *Marchantia polymorpha* is mediated by evolutionarily conserved promoter elements. *Physiol.*
904 *Plantarum* 156, 407-420.
- 905 55. Gehrman, K. (1909). Zur befruchtungsphysiologie von *Marchantia polymorpha* L. *Berichte*
906 *der Deutschen Botanischen Gesellschaft* 27, 341-348.
- 907 56. Goebel, K. (1898). *Organographie der pflanzen, insbesondere der archegoniaten und*
908 *samenpflanzen*, Volume 2, 3rd Edition, (Jena: Gustav Fischer).
- 909 57. Proctor, M.C.F. (2008). The physiological basis of bryophyte production. *Bot. J. Linn. Soc.*
910 104, 61-77.
- 911 58. Asakawa, Y. (2008). Liverworts-potential source of medicinal compounds. *Curr. Pharm. Des.*
912 14, 3067-3088.
- 913 59. Jia, Q., Li, G., Kollner, T.G., Fu, J., Chen, X., Xiong, W., Crandall-Stotler, B.J., Bowman, J.L.,
914 Weston, D.J., Zhang, Y., et al. (2016). Microbial-type terpene synthase genes occur widely in
915 nonseed land plants, but not in seed plants. *Proc. Natl. Acad. Sci. USA* 113, 12328-12333.
- 916 60. Kanazawa, T., Era, A., Minamino, N., Shikano, Y., Fujimoto, M., Uemura, T., Nishihama, R.,
917 Yamato, K.T., Ishizaki, K., Nishiyama, T., et al. (2016). SNARE molecules in *Marchantia*
918 *polymorpha*: unique and conserved features of the membrane fusion machinery. *Plant Cell*
919 *Physiol.* 57, 307-324.
- 920 61. Hofer, J., Turner, L., Moreau, C., Ambrose, M., Isaac, P., Butcher, S., Weller, J., Dupin, A.,
921 Dalmais, M., Le Signor, C., et al. (2009). *Tendrill-less* regulates tendril formation in pea leaves.
922 *Plant Cell* 21, 420-428.
- 923 62. Vlad, D., Kierzkowski, D., Rast, M.I., Vuolo, F., Dello Ioio, R., Galinha, C., Gan, X.,
924 Hajheidari, M., Hay, A., Smith, R.S., et al. (2014). Leaf shape evolution through duplication,
925 regulatory diversification, and loss of a homeobox gene. *Science* 343, 780-783.

- 926 63. Perroud, P.F., Haas, F.B., Hiss, M., Ullrich, K.K., Alboresi, A., Amirebrahimi, M., Barry, K.,
927 Bassi, R., Bonhomme, S., Chen, H., et al. (2018). The *Physcomitrella patens* gene atlas project:
928 large-scale RNA-seq based expression data. *Plant J.*
- 929 64. Zwickel, W. (1932). Studien über die ocellen der lebermoose. Beihefte zum Botanischen
930 Centralblatt 49, 569-648.
- 931 65. Garjeanne, A.J.M. (1903). Die ölkörper der Jungermanniales. *Flora* 92, 457-482.
- 932 66. Labandeira, C. (2007). The origin of herbivory on land: initial patterns of plant tissue
933 consumption by arthropods. *Insect Sci.* 14, 259-275.
- 934 67. Peñuelas, M., Monte, I., Schweizer, F., Vallat, A., Reymond, P., García-Casado, G., Franco-
935 Zorrilla, J.M., and Solano, R. (2019). Jasmonate-related MYC transcription factors are
936 functionally conserved in *Marchantia polymorpha*. *Plant Cell* 31, 2491-2509.
- 937 68. Lange, B.M. (2015). The evolution of plant secretory structures and emergence of terpenoid
938 chemical diversity. *Annu. Rev. Plant. Biol.* 66, 139-159.
- 939 69. Bowman, J.L. (2016). A brief history of *Marchantia* from Greece to genomics. *Plant Cell*
940 *Physiol.* 57, 210-229.
- 941 70. Rimington, W.R., Pressel, S., Duckett, J.G., Field, K.J., Read, D.J., and Bidartondo, M.I.
942 (2018). Ancient plants with ancient fungi: liverworts associate with early-diverging arbuscular
943 mycorrhizal fungi. *Proc. Biol. Sci.* 285.
- 944 71. Carella, P., Gogleva, A., Hoey, D.J., Bridgen, A.J., Stolze, S.C., Nakagami, H., and Schornack,
945 S. (2019). Conserved biochemical defenses underpin host responses to oomycete infection in
946 an early-divergent land plant lineage. *Curr. Biol.* 29, 2282-2294 e2285.
- 947 72. Flores-Sandoval, E., Eklund, D.M., and Bowman, J.L. (2015). A simple auxin transcriptional
948 response system regulates multiple morphogenetic processes in the liverwort *Marchantia*
949 *polymorpha*. *PLoS Genet* 11, e1005207.
- 950 73. Busch, A., Deckena, M., Almeida-Trapp, M., Kopischke, S., Kock, C., Schussler, E., Tsiantis,
951 M., Mithofer, A., and Zachgo, S. (2019). MpTCP1 controls cell proliferation and redox
952 processes in *Marchantia polymorpha*. *New Phytol.* 224, 1627-1641.
- 953 74. Ishizaki, K., Nishihama, R., Yamato, K.T., and Kohchi, T. (2016). Molecular genetic tools and
954 techniques for *Marchantia polymorpha* research. *Plant Cell Physiol.* 57, 262-270.
- 955 75. Ishizaki, K., Chiyoda, S., Yamato, K.T., and Kohchi, T. (2008). Agrobacterium-mediated
956 transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for
957 plant biology. *Plant Cell Physiol.* 49, 1084-1091.
- 958 76. Kubota, A., Ishizaki, K., Hosaka, M., and Kohchi, T. (2013). Efficient Agrobacterium-mediated
959 transformation of the liverwort *Marchantia polymorpha* using regenerating thalli. *Biosci.*
960 *Biotechnol. Biochem.* 77, 167-172.
- 961 77. Sugano, S.S., Shirakawa, M., Takagi, J., Matsuda, Y., Shimada, T., Hara-Nishimura, I., and
962 Kohchi, T. (2014). CRISPR/Cas9-mediated targeted mutagenesis in the liverwort *Marchantia*
963 *polymorpha* L. *Plant Cell Physiol.* 55, 475-481.
- 964 78. Sugano, S.S., Nishihama, R., Shirakawa, M., Takagi, J., Matsuda, Y., Ishida, S., Shimada, T.,
965 Hara-Nishimura, I., Osakabe, K., and Kohchi, T. (2018). Efficient CRISPR/Cas9-based genome
966 editing and its application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS One*
967 13, e0205117.
- 968 79. Hisanaga, T., Okahashi, K., Yamaoka, S., Kajiwar, T., Nishihama, R., Shimamura, M.,
969 Yamato, K.T., Bowman, J.L., Kohchi, T., and Nakajima, K. (2019). A cis-acting bidirectional
970 transcription switch controls sexual dimorphism in the liverwort. *EMBO J.* 38.

971 80. Ishizaki, K., Nishihama, R., Ueda, M., Inoue, K., Ishida, S., Nishimura, Y., Shikanai, T., and
972 Kohchi, T. (2015). Development of Gateway binary vector series with four different selection
973 markers for the liverwort *Marchantia polymorpha*. PLoS One 10, e0138876.

974 81. Hellens, R.P., Edwards, E.A., Leyland, N.R., Bean, S., and Mullineaux, P.M. (2000). pGreen:
975 a versatile and flexible binary Ti vector for Agrobacterium-mediated plant transformation. Plant
976 Mol. Biol. 42, 819-832.

977 82. Eshed, Y., Baum, S.F., and Bowman, J.L. (1999). Distinct mechanisms promote polarity
978 establishment in carpels of Arabidopsis. Cell 99, 199-209.

979 83. Barbier de Reuille, P., Routier-Kierzkowska, A.L., Kierzkowski, D., Bassel, G.W., Schupbach,
980 T., Tauriello, G., Bajpai, N., Strauss, S., Weber, A., Kiss, A., et al. (2015). MorphoGraphX: a
981 platform for quantifying morphogenesis in 4D. Elife 4, 05864.

982 84. Re, D.A., Capella, M., Bonaventure, G., and Chan, R.L. (2014). Arabidopsis *AtHB7* and
983 *AtHB12* evolved divergently to fine tune processes associated with growth and responses to
984 water stress. BMC Plant Biol. 14, 150.

985 85. Nakazaki, A., Yamada, K., Kunieda, T., Sugiyama, R., Hirai, M.Y., Tamura, K., Hara-
986 Nishimura, I., and Shimada, T. (2019). Leaf endoplasmic reticulum bodies identified in
987 Arabidopsis rosette leaves are involved in defense against herbivory. Plant Physiol. 179, 1515-
988 1524

989 86. Hockett, K.L., and Baltrus, D.A. (2017). Use of the soft-agar overlay technique to screen for
990 bacterially produced inhibitory compounds. J. Vis. Exp. 119, 55064.

991 87. Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2:
992 accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.
993 Genome Biol. 14, R36.

994 88. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and
995 Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. 29, 24-26.

996 89. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-
997 throughput sequencing data. Bioinformatics 31, 166-169.

998 90. Afgan, E., Sloggett, C., Goonasekera, N., Makunin, I., Benson, D., Crowe, M., Gladman, S.,
999 Kowsar, Y., Pheasant, M., Horst, R., et al. (2015). Genomics virtual laboratory: a practical
1000 bioinformatics workbench for the cloud. PLoS One 10, e0140829.

1001 91. Flores-Sandoval, E., Eklund, D.M., Hong, S.F., Alvarez, J.P., Fisher, T.J., Lampugnani, E.R.,
1002 Golz, J.F., Vazquez-Lobo, A., Dierschke, T., Lin, S.S., et al. (2018). Class C ARFs evolved
1003 before the origin of land plants and antagonize differentiation and developmental transitions in
1004 *Marchantia polymorpha*. New Phytol. 218, 1612-1630.

1005 92. Montgomery, S.A., Tanizawa, Y., Galik, B., Wang, N., Ito, T., Mochizuki, T., Akimcheva, S.,
1006 Bowman, J.L., Cognat, V., Marechal-Drouard, L., et al. (2020). Chromatin organization in early
1007 land plants reveals an ancestral association between H3K27me3, transposons, and constitutive
1008 heterochromatin. Curr Biol.

1009 93. Davies, K.M., Jibrán, R., Zhou, Y., Albert, N.W., Brummell, D.A., Jordan, B.R., Bowman, J.L.,
1010 and Schwinn, K.E. (2020). The evolution of flavonoid biosynthesis: a bryophyte perspective.
1011 Front Plant Sci 11.

1012 94. Adams, R.P. (2009). Identification of essential oil components by gas chromatography/mass
1013 spectroscopy, (IL, USA.: Carol Stream, Ill. : Allured Pub. Corp.).

1014

Supplementary Methods and Figures

Supplementary Methods

Transient transformation of tobacco leaves

Nicotiana benthamiana leaves were syringe-infiltrated with a bacterial suspension of *Agrobacterium* (LBA4404) harboring the construct *pro35S::GFP-MpC1HDZ* or water (mock) following a previously established protocol [1]. The abaxial epidermis of leaves was imaged using confocal microscopy (Leica TCS SP8 Compact, Leica Microsystems, Wetzlar, Alemania). For GFP channel an excitation at 488nm, and emission at 498-510nm; for chloroplast autofluorescence channel we used the same excitation wave length and emission at 595-669nm. Images were acquired and processed using Leica Application Suite (LAS-AF, v.2.7.3) and ImageJ software (National Institute of Health, <https://imagej.nih.gov/ij/>) was used for generating merged pictures.

RT-qPCR analysis

Total RNA was purified from 3-week-old *M. polymorpha* plants using Trizol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. 1µg of RNA was reverse-transcribed using oligo(dT)18 and MMLV reverse transcriptase II (Promega, Madison, WA, USA). Quantitative real-time PCR (qPCR) was performed using a StepOnePlus Real-Time PCR System (Applied Biosystems); each reaction contained a final volume of 20 µL that included 2 µL of SyBr green (4×), 8 pmol of each primer, 2 mM MgCl₂, 10 µL of a 1/15 dilution of the RT reaction and 0.1 µL of Taq Polymerase (Invitrogen). Thermocycler parameters were as follows: 94°C 12 sec, 60°C 12 sec, 72°C 12 sec. Fluorescence was quantified over 40 cycles at 72°C. Specific primers for each gene were designed and are listed in Table S5. The mRNA levels were quantified by normalizing their levels to the levels of the ACTIN 7 [2] using the Ct method. All of the reactions were performed with three biological replicates. Each replicate was obtained by pooling tissue from 3 to 4 individual plants.

Literature

1. de Felippes, F.F., and Weigel, D. (2010). Transient assays for the analysis of miRNA processing and function. *Methods Mol. Biol.* 592, 255-264.
2. Saint-Marcoux, D., Proust, H., Dolan, L., and Langdale, J.A. (2015). Identification of reference genes for real-time quantitative PCR experiments in the liverwort *Marchantia polymorpha*. *PLoS One* 10, e0118678.

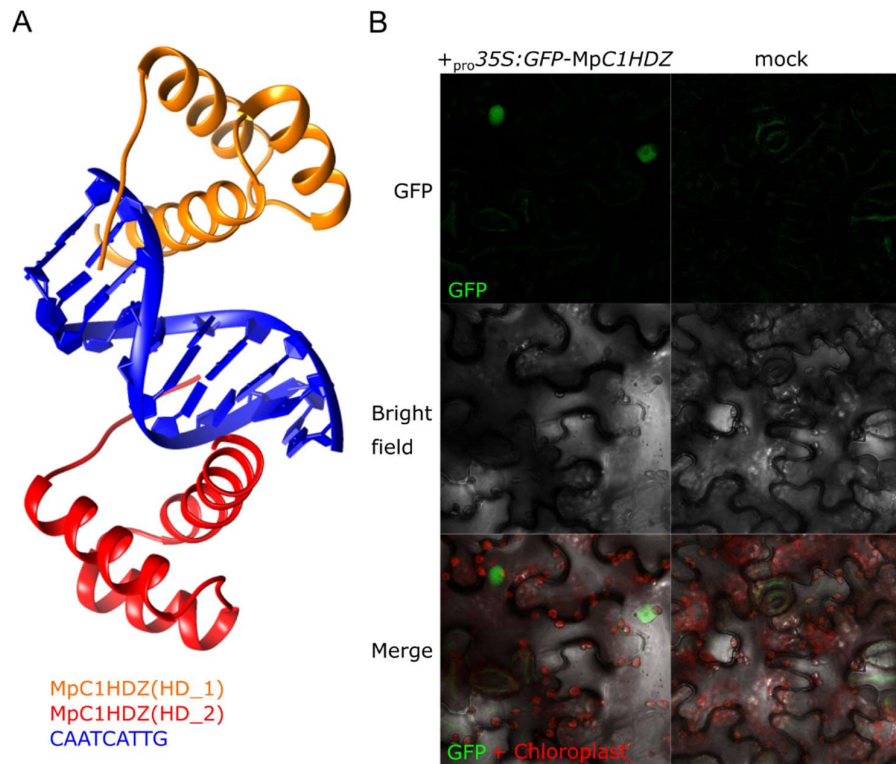


Figure S1. MpC1HDZ resembles structural features of angiosperms C1HDZ orthologs. (A) Ribbon representation of the 3D model of two MpC1HDZ DNA binding domains interacting with the pseudo-palindromic sequence CAATCATTG. The structure was determined using Phyre2 and based on a *Drosophila* homeobox protein (PDB: 1JGG). (B) Subcellular localization of MpC1HDZ in *Nicotiana benthamiana* leaves infiltrated with a bacterial suspension of *A. tumefaciens* harboring the construct *pro35S:GFP-MpC1HDZ* or mock (water). The abaxial epidermis of rosette leaves was imaged using confocal microscopy. The photos represent the GFP specific channel, bright-field, and an overlay of both (Merge).

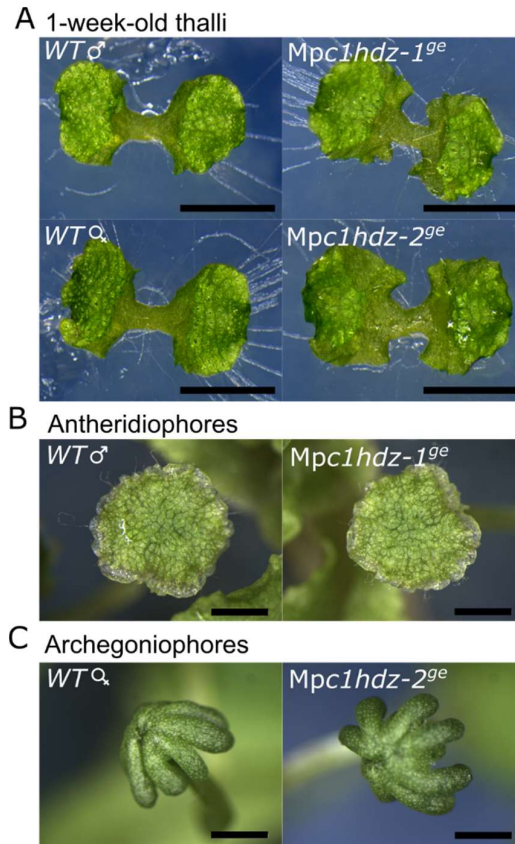


Figure S2. Plant morphology of wild-type and mutant alleles. Thallus of wild-type and *Mpc1hdz* plants 7 days after germination (DAG). Wild-type and *Mpc1hdz* gametangiophores; male, antheridiophore; female, archegoniophore. Scale bars (2mm).

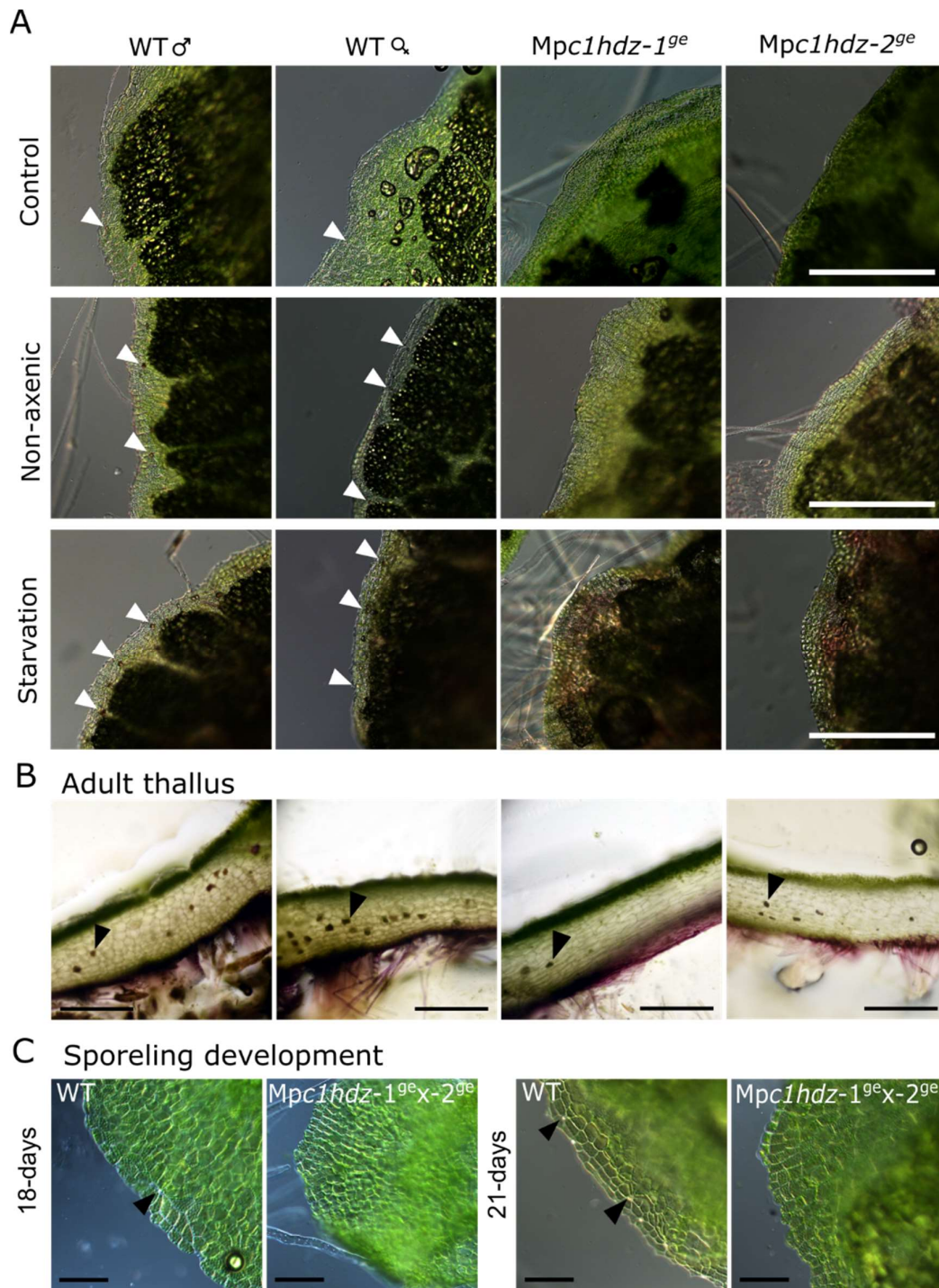


Figure S3. Plant growth responses upon abiotic stress. (A) Representative compound microscopy pictures of 3-week-old plants thallus grown in control conditions for one week and then transferred again to control conditions, or starvation conditions (1/100 Gamborg's B5 medium) or non-axenic (rockwool as substrate). Arrows point to oil bodies. Scale bars (0.5mm). (B) Transverse sections of 30-day-old thallus grown in non-axenic conditions. (C) Representative pictures of wild-type and *Mpc1hdz* sporelings. Black arrow indicates oil bodies. Scale bars (0.1mm).

WT +
proMpc1hdz:MpC1HDZ_{mut}:Citrine

WT +
pro35S:MpC1HDZ_{mut}:Citrine

WT +
proEF1:MpC1HDZ_{mut}

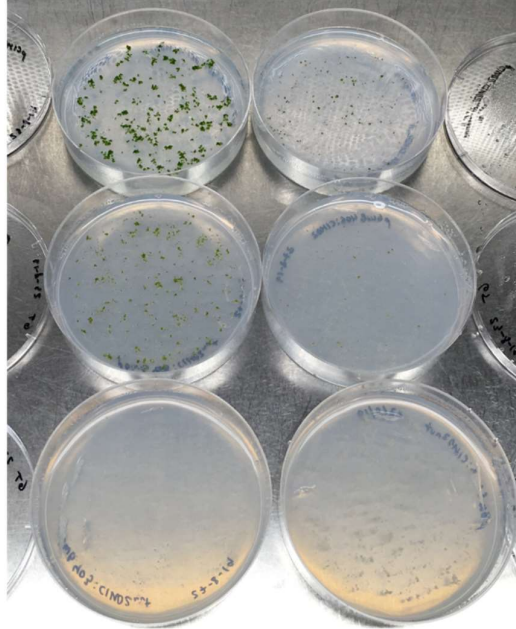


Figure S4. Constitutive overexpression of *MpC1HDZ* may be lethal. Representative pictures of wild type (*WT*) spores transformed with a vector constitutively overexpressing *MpC1HDZ*. After 10 day of selection on antibiotic plates, most plants did not survive compared to a control transformation. Two biological replicates are shown.

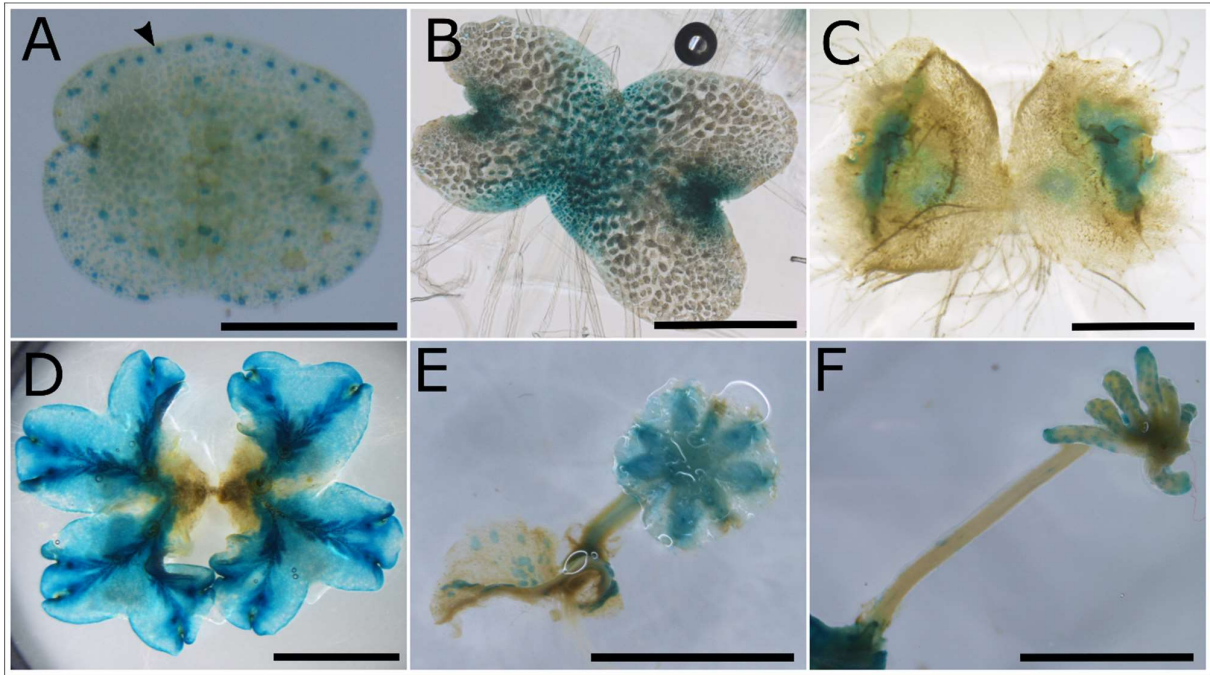


Figure S5. *MpCIHDZ* promoter activity in BoGa ecotype. *MpCIHDZ* expression pattern through different life stages visualized by β -glucuronidase staining in *proMpCIHDZ*^{3kb}:*GUS* lines (A – F). Vegetative propagule, gemma (A), 3 day-old gemmaling (B), one week old gemmaling (C), three weeks old mature thallus (D), male sexual structure, antheridiophore (E) and female sexual structure, archegoniophore (F).

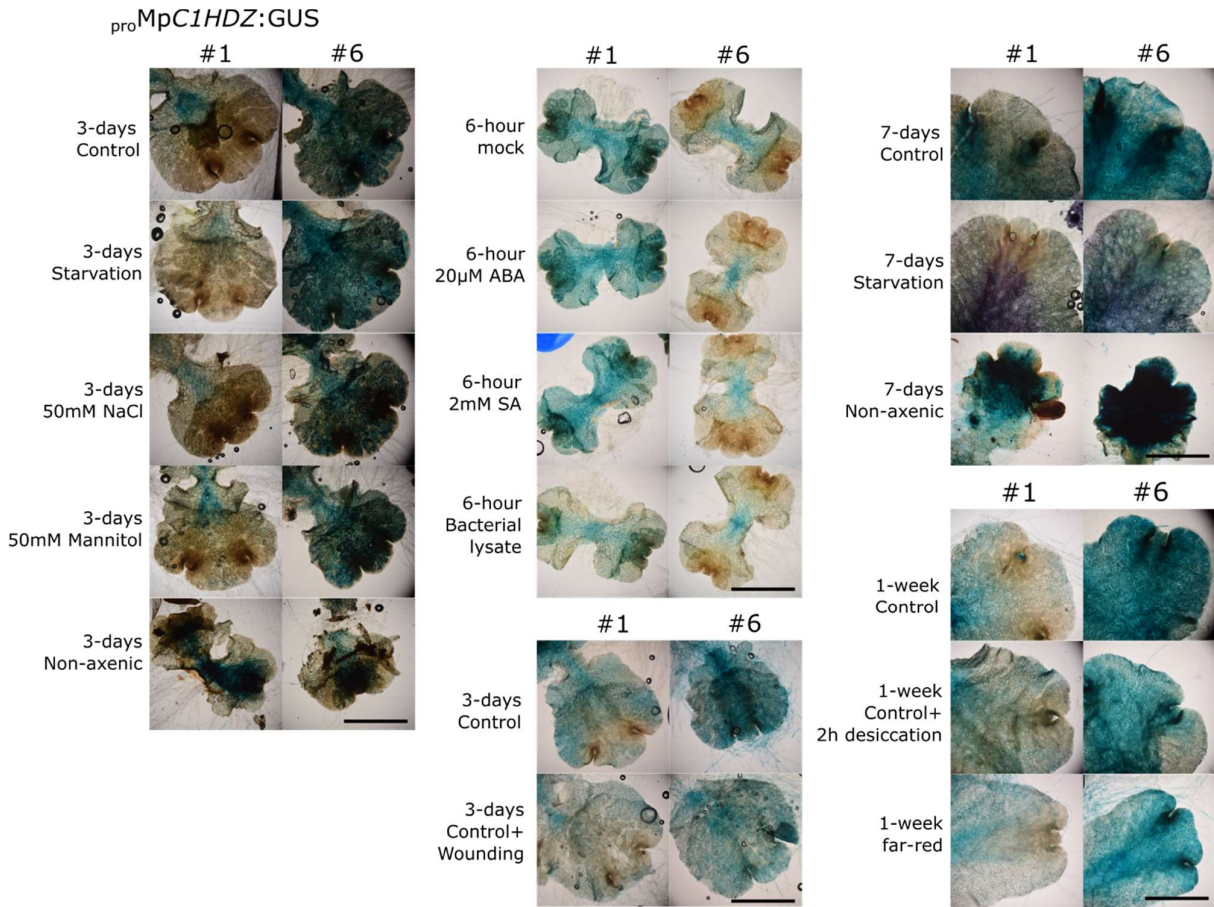


Figure S6. *MpC1HDZ* promoter activity upon abiotic stress treatments. GUS staining of *proMpC1HDZ^{4.6kb}::GUS* transcriptional reporter under different treatments. Plants were grown for one week and then transferred to new plates containing different media and treatments as indicated in the left. Scale bars represent 0.5mm.

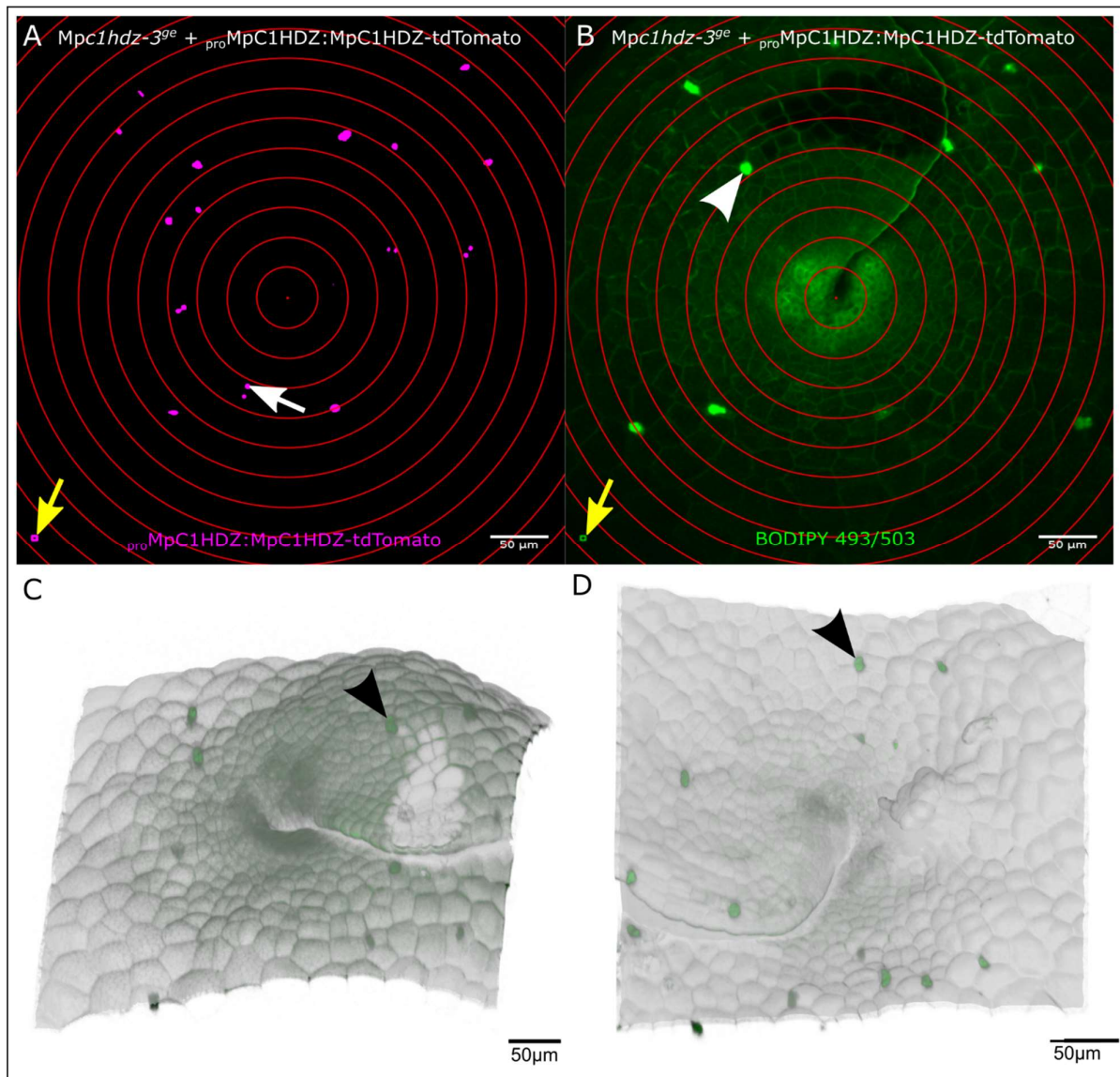


Figure S7. Measurement of distance from apical notch to nuclei highly expressing MpC1HDZ and oil bodies. Apical notch area of *Mpc1hdz-3* gemmaling complemented with *proMpC1HDZ^{3kb}:MpC1HDZ-tdTomato* and stained with BODIPY 493/503 dye. (A, B) 2D projection of a representative confocal z-stack from Fig. 2G with concentric circles used for quantification in Fig. 2H. MpC1HDZ-tdTomato channel with nuclei above the threshold (A) and BODIPY 493/503 channel with oil bodies (B). Concentric circles are centered in the middle of the apical notch and are interspaced by 25 μm. White arrow points to a bright nucleus, arrowhead points to an oil body, and yellow arrows point to the pixels used for calibration of “sum slices” z-stack projection. (C, D) 3D representation of apical notch surface from two replicates with BODIPY signal 493/503 shown in green. Arrowheads point to oil bodies. The sample in C corresponds to A and B. Note that the circles

in A and B represent approximate measure because of variability in apical notch area 3D morphology (C, D). Scale bars represent 50 μm .

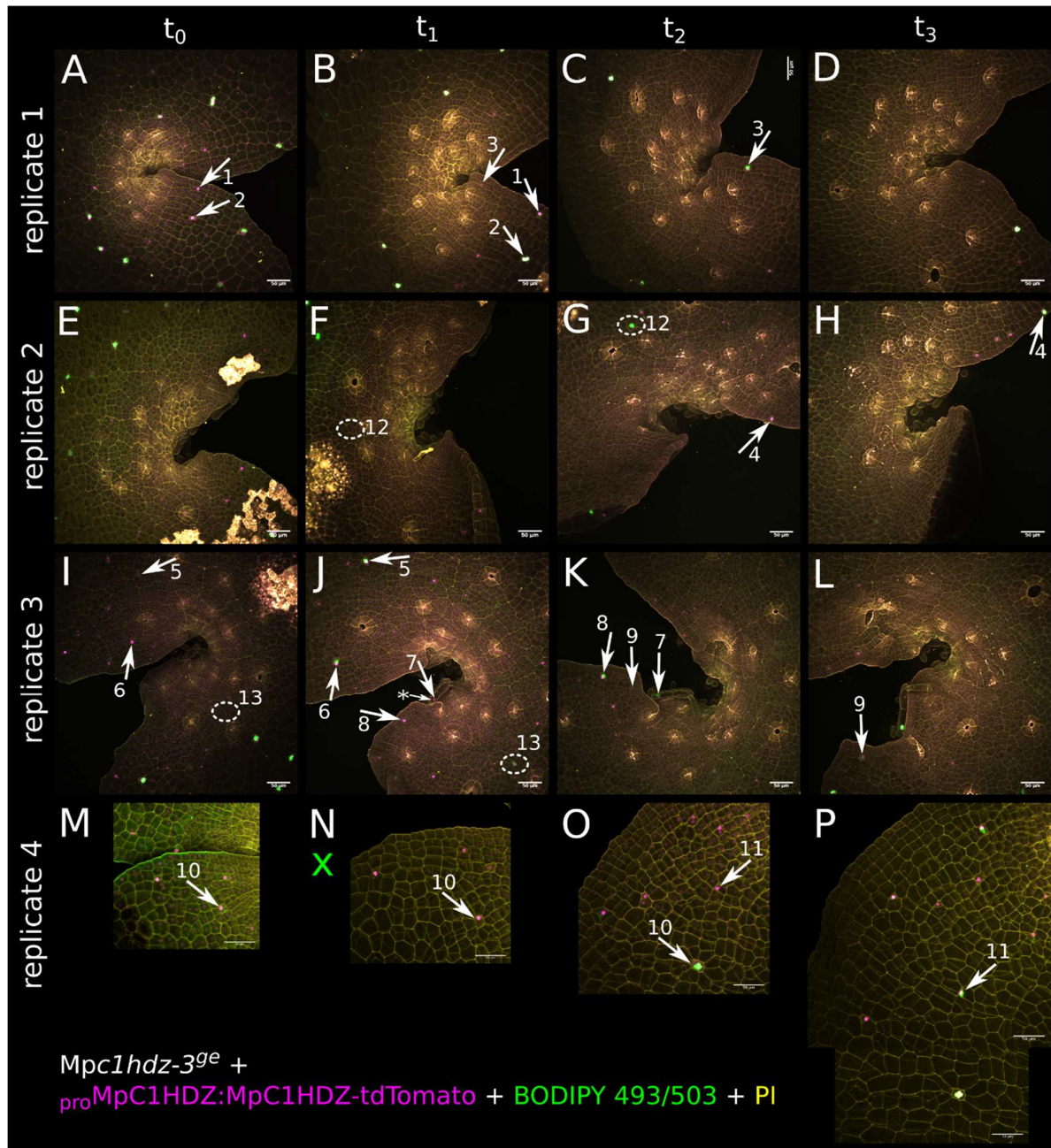


Figure S8. Overview of oil body cell tracking time lapse. Time series of oil body cell differentiation in apical notch area of 4 day-old *Mpc1hdz-3* gemmalings complemented with *proMpC1HDZ3kb:MpC1HDZ-tdTomato* (magenta) and stained with BODIPY 493/503 (green) and PI (yellow) directly before CLSM imaging. The same area was followed for 48 h in 12 h intervals (t_0 – t_1 , shown from left to right). Standard deviation projections of confocal z-stacks of four replicates are shown: replicate 1 (A – D), replicate 2 (E – H), replicate 3 (I – L) and replicate 4 (M – P). All newly-formed oil body cells are equivalently numbered and marked before and after oil body formation event. Newly-formed oil body cells that had exceptionally high *MpC1HDZ-tdTomato*

expression compared to surrounding cells in previous time point are indicated with arrows (numbers 1 – 11). Newly-formed oil body cells that did not have outstandingly bright MpC1HDZ-tdTomato signal in previous time point are marked with dashed circles (numbers 12 and 13). The latter observations may suggest that the time interval used for observation is longer than the minimal period of high MpC1HDZ expression needed for oil body formation. Alternatively, high MpC1HDZ expression may not be necessary for oil body formation in a small proportion of cells, which would also explain presence of a few oil bodies in *Mpc1hdz* mutants (Figure 1, Figure S3). Asterisk and a small arrow mark a shadow observed in cell 7 at the spot stained with BODIPY in the next time point. This may indicate that this cell was developing an oil body, but did not yet accumulate the compounds stainable with BODIPY by the time of observation. This was also the only observed oil body formation in a large cell that most likely belongs to a ventral scale. One case where BODIPY was not applied is marked with green X. Signal intensities between the images are not comparable due to varying z-stack sizes. Scale bars represent 50 μm .

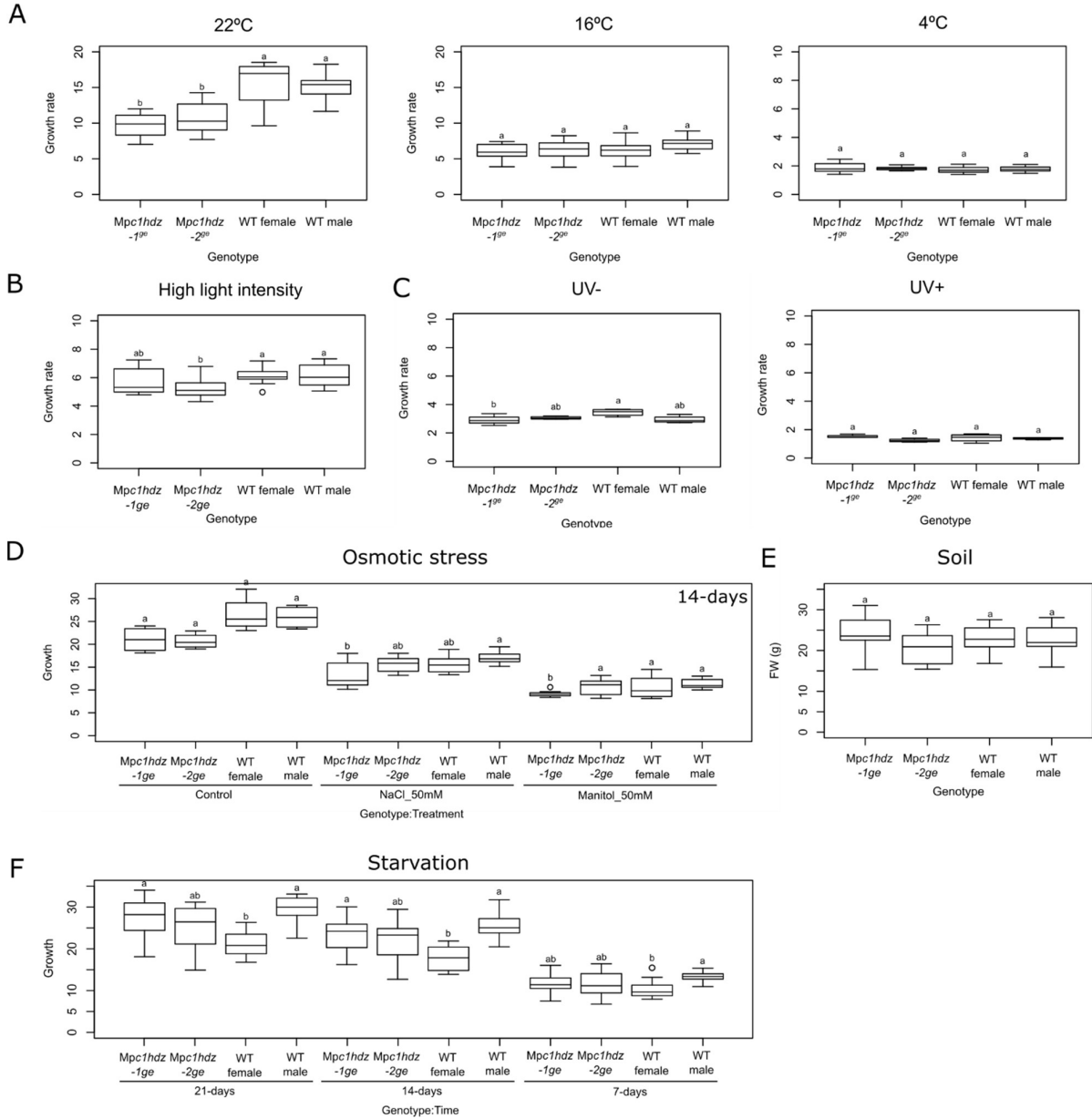


Figure S9. Thalli growth rate of *Mpc1hdz* under different abiotic stress treatments. (A-F) Phenotypic effect of abiotic stress treatments on wild-type and *Mpc1hdz* plants. Plants were grown axenically on agar plates with 1/2 strength Gamborg B5 media for 1-week and then transferred to a new plate. Growth was measured as lamina exposed area compared to the area after 1 week of treatment or different time-points. Cold treatment was performed at 4°C; high light intensity (HLI) was performed using 1000 $\mu\text{mol m}^{-2}\text{s}^{-1}$ LED lights; UV treatment was performed in a laminar flow bench UV-C light for 1 hour and recovered for an additional week under control conditions. For osmotic stress, media was supplemented with 50 mM NaCl or mannitol, grown for 2-weeks (in Figure 1 it is shown after 1-week). For starvation treatments, plants were placed in 1/100 Gamborg B5 media.

For non-axenic growth, plants were grown on pots in a growth chamber using rockwool as a substrate. Statistical difference was tested using ANOVA followed by Tukey HSD p .value < 0.05; letters indicate statistically significant groups. Representative pictures of each treatment are shown.

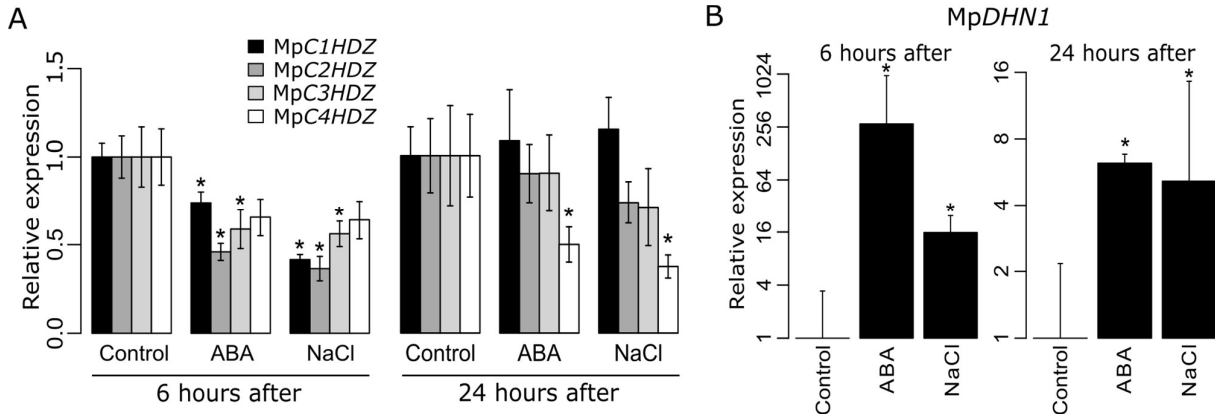


Figure S10. Expression of *HDZ* genes under ABA and abiotic stress treatments. (A) RT-qPCR analysis of all *HDZ* genes in *M. polymorpha* under ABA (20 uM) and NaCl (200 mM) treatments. RNA samples were taken after 6 and 24 h. (B) The abiotic stress marker gene (*MpDHN1*) was used as a treatment control. Statistical differences are marked as asterisks and computed pairwise *t*-test (p .value < 0.05) comparing each treatment to the corresponding control.

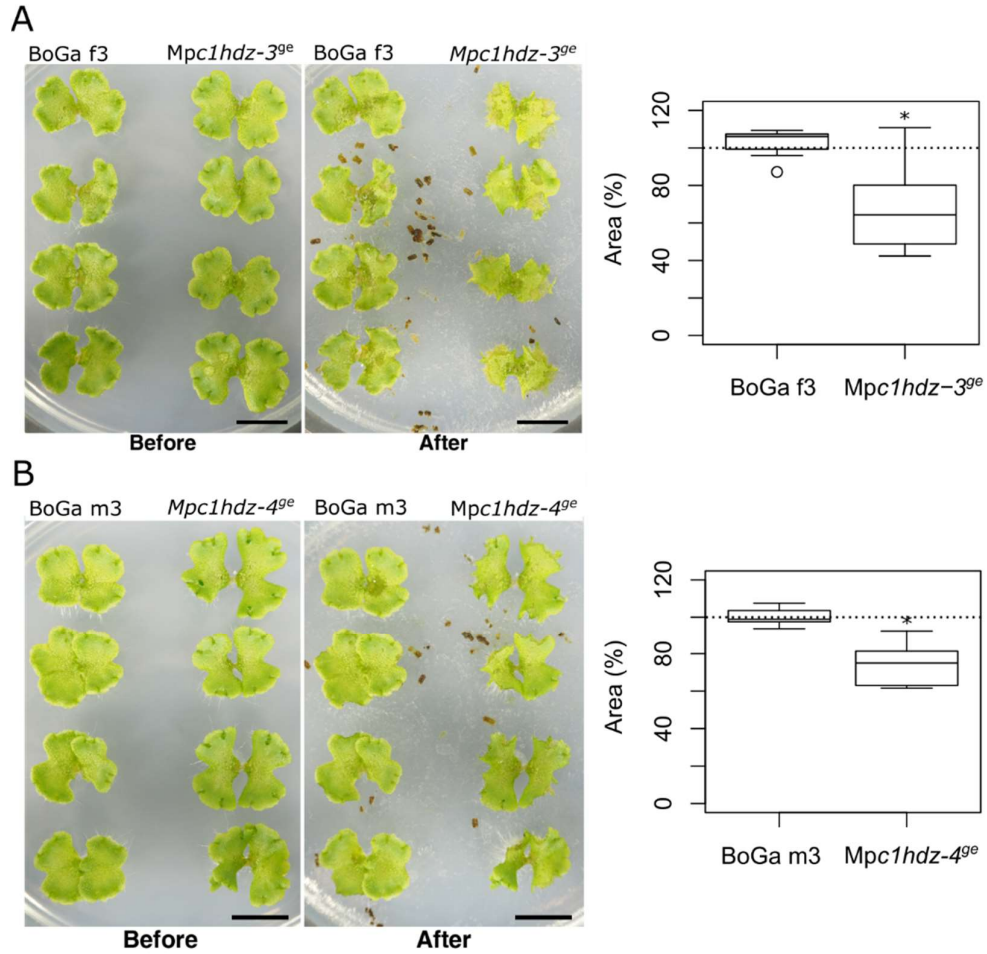


Figure S11. Pill bug feeding assay on mutant alleles of the BoGa ecotype. Starved pill bugs and 13-day-old thalli (Before) were co-cultivated for 24 hours (After). Female wild-type (BOGA f3) and *Mpc1hdz-3^{ge}* mutant thalli or male wild-type (BOGA m3) and *Mpc1hdz-4^{ge}* mutant thalli were used in the experiments. Bars = 1 cm. Thallus area change indicates the difference between thallus area (mm²) of “After” and that of “Before.” Thallus change ratio indicates the ratio of thallus area of “After” to that of “Before” (n = 12 thalli for each genotype). Error bars indicate the mean $\pm$ SD. Statistical analysis between wild-type and mutant plants was conducted by a two-tailed Welch’s *t*-test.

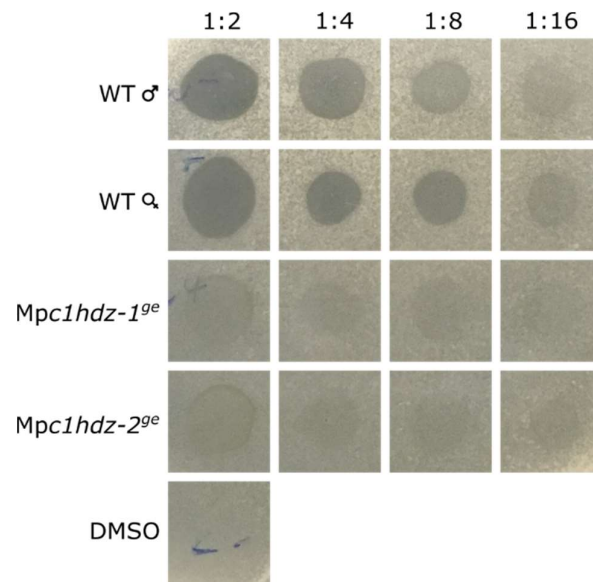


Figure S12. Dose-dependence effect of plant extract on the growth of *Bacillus subtilis*. Serial dilution of *Bacillus subtilis* overlay assay using ethanol extracts in DMSO from wild-type and Mpc1hdz mutant thalli.

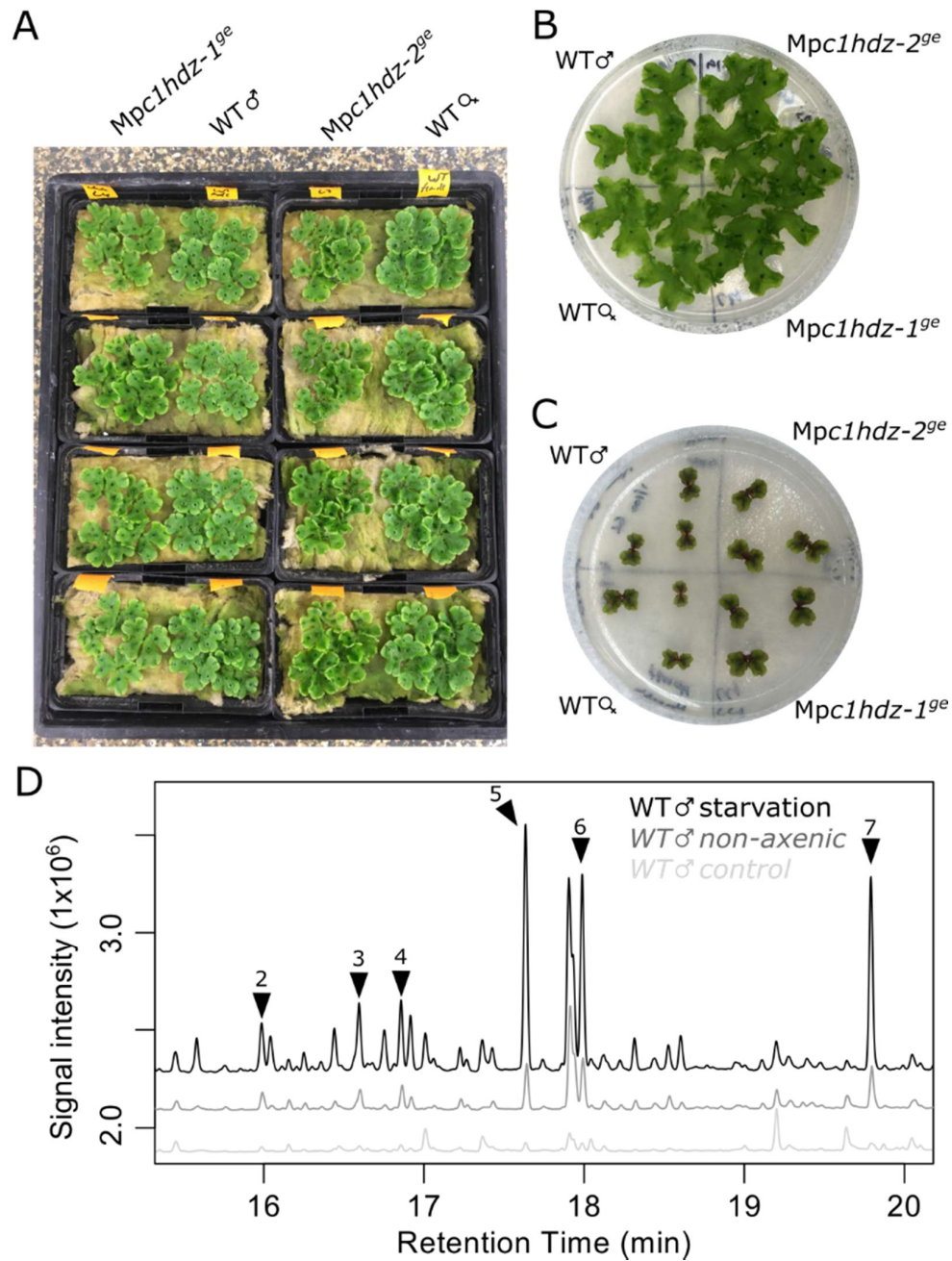


Figure S13. Representative figures of plants after treatments. (A) 3-week old wild-type and *Mpc1hdz* plants growing in non-axenic conditions (rockwool) (B) 3-week old wild-type and *Mpc1hdz* plants growing in axenic conditions 1/2 Gamborg B5 (control). (C) 3-week old wild-type and *Mpc1hdz* plants growing in starvation treatments 1/100 Gamborg B5.

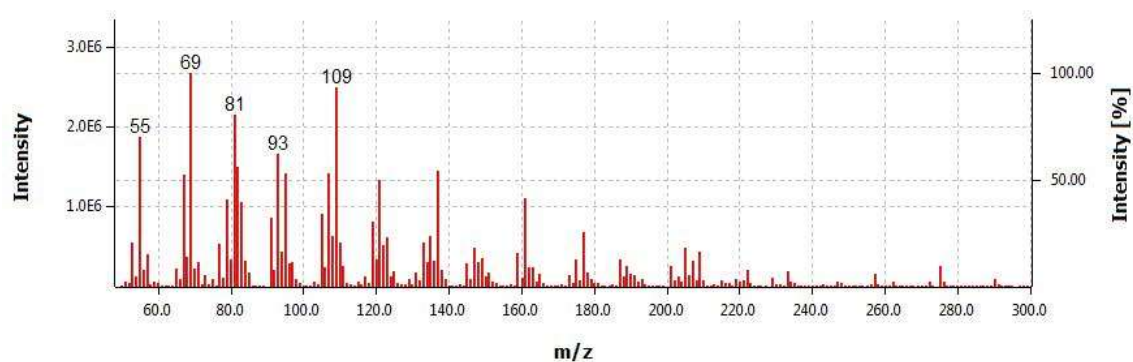


Figure S14. Mass spectrum of unidentified compound U13. Mass spectrum of peak 10 corresponding to the unidentified compound U13. Based on MS profiles of NIST library, we tentatively propose it is a terpenoid.