

Title

Stomatal regulators are co-opted for seta development in the astomatous liverwort *Marchantia polymorpha*

Author list

Kenta C. Moriya¹, Makoto Shirakawa², Jeanne Loue-Manifel^{3,4}, Yoriko Matsuda⁵, Yen-Ting Lu^{2,4}, Kentaro Tamura⁶, Yoshito Oka¹, Tomonao Matsushita¹, Ikuko Hara-Nishimura⁷, Gwyneth Ingram³, Ryuichi Nishihama^{5,8}, Justin Goodrich⁴, Takayuki Kohchi⁵, Tomoo Shimada^{1,*}

Affiliations

¹Graduate School of Science, Kyoto University, Kyoto 606-8502, Japan

²Graduate School of Biological Sciences, Nara Institute of Science and Technology (NAIST), Ikoma 630-0192, Japan

³Laboratoire Reproduction et Développement des Plantes, ENS de Lyon, CNRS, INRAE, UCB Lyon 1, Lyon 69342, France

⁴Institute of Molecular Plant Sciences, University of Edinburgh, Daniel Rutherford Building, Max Born Crescent, Edinburgh EH9 3BF, UK

⁵Graduate School of Biostudies, Kyoto University, Kyoto 606-8502, Japan

⁶School of Food and Nutritional Sciences, University of Shizuoka, Shizuoka 422-8526, Japan

⁷Faculty of Science and Engineering, Konan University, Kobe 658-8501, Japan

⁸Department of Applied Biological Science, Faculty of Science and Technology, Tokyo University of Science, Noda 278-8510, Japan

*Correspondence: tshimada@gr.bot.kyoto-u.ac.jp

Abstract

The evolution of special types of cells requires the acquisition of new gene regulatory networks controlled by transcription factors (TFs). In stomatous plants, a TF module formed by subfamilies Ia and IIIb basic helix-loop-helix TFs (Ia-IIIb bHLH) regulates stomatal formation; however, how this module evolved during land plant diversification remains unclear. Here, we show that, in the astomatous liverwort *Marchantia polymorpha*, a Ia-IIIb bHLH module regulates the development of a unique sporophyte tissue, the seta, which is found in mosses and liverworts. The sole Ia bHLH gene, MpSETA, and a IIIb bHLH gene, MpICE2, regulate the cell division and/or differentiation of seta lineage cells. MpSETA can partially replace the stomatal function of Ia bHLH TFs in *Arabidopsis thaliana*, suggesting that a common regulatory mechanism underlies the setal and stomatal formation. Our findings reveal the co-option of a Ia-IIIb bHLH TF module for regulating cell fate determination and/or cell division of distinct types of cells during land plant evolution.

Main text

Land plants developed unique types of cells and tissues to adapt to the terrestrial environment during evolution^{1,2}. The acquisition of novel cells or tissues leading to complex body plans is related to the diversification of transcription factors (TFs)^{3,4}. Basic helix-loop-helix (bHLH) TFs represent a TF superfamily that plays key roles in cell fate determination and cell division during eukaryote development. In land plants, the number of genes encoding bHLH TFs has increased compared with those of chlorophyte and charophyte algae, suggesting that bHLH TFs are involved in terrestrial adaptation⁵. For example, the acquisition of stomata, a special tissue for gas exchange on the epidermis, is an important adaptation of plants to terrestrial environments; previous studies have revealed that stomatal formation is regulated by bHLH TFs as master TFs⁶⁻⁹. In *Arabidopsis thaliana*, three TFs belonging to subfamily Ia (SPEECHLESS [SPCH], MUTE, and FAMA)⁶⁻⁸ form heterodimers with subfamily IIIb TFs (ICE1, also known as SCREAM [SCRM], and ICE2, also known as SCRM2) to promote the stomatal formation by regulating downstream gene expression⁹. The molecular mechanism of stomatal formation by Ia and IIIb bHLH TFs is conserved in the moss *Physcomitrium patens*, in which stomata play an important role in spore dispersal by promoting dehydration and dehiscence of the sporangium^{10,11}.

Recent studies have revealed that a gene encoding a Ia bHLH is present in the genome of the astomatous liverwort *Marchantia polymorpha*^{4,12-14}. Despite the importance of Ia bHLH in stomatal development, its function in plants without stomata remains unexplored. Here, we report that the Ia bHLH protein, designated as MpSETA, is a master regulator of the formation of the seta, which is a diploid tissue involved in long-distance spore dispersal in *M. polymorpha*. Furthermore, we show that Ia and IIIb bHLH positively regulate setal formation by heterodimerization, similar to their role in stomatal formation in other land plants. This outcome advances our understanding of the mechanisms of the evolution of plant tissue formation while providing new insights into the co-option of gene expression regulatory networks (GRNs).

Results

MpSETA is the sole Ia bHLH protein in *M. polymorpha*.

To identify Ia bHLH coding genes in *M. polymorpha*, we constructed a phylogenetic tree of Ia bHLH proteins from various plant species using a bHLH domain and C-terminal conserved domain called SMF (also known as the ACT-like domain)^{15,16} (Fig. 1a and Extended Data Fig. 1a). Our phylogenetic analysis suggested that MpBHLH35 (Mp2g04160) is the sole Ia bHLH in *M. polymorpha* (Fig. 1a). We named this gene MpSETA based on its specific expression in the seta tissue of the sporophyte (see below). Multiple alignments of the bHLH proteins revealed that the amino acid residues predicted to be important for E-box (CANNTG) binding and bHLH dimerization are highly conserved in MpSETA, although the amino acid sequence of the bHLH domain of MpSETA is relatively divergent compared to other Ia bHLH (Extended Data Fig. 1b). Additionally, we found partial sequences of two MpSETA related genes (LcSETA1 and LcSETA2) in the genome of the Marchantiidae liverwort *Lunularia cruciata*¹⁷, although there is no evidence that these putative MpSETA-like genes are expressed and are functional in this species (Fig. 1a and Extended Data Fig. 1b-c). The amino acid sequences of the bHLH and SMF domains are well conserved between MpSETA and LcSETA1.

Thus, we can conclude that Ia bHLH genes are conserved in the genome of Marchantiidae liverworts.

Because the amino acid sequence of MpSETA is divergent, whether MpSETA shares similar properties with the other Ia bHLH proteins is unclear. Therefore, we investigated whether MpSETA can act as a stomatal regulator replacing AtSPCH, AtMUTE, or AtFAMA. In this context, MpSETA was expressed under the native promoters of AtSPCH, AtMUTE, and AtFAMA in *spch-3*, *mute-2*, and *fama-1* backgrounds, respectively (Fig. 1b and Extended Data Fig. 2a,b). Even though *mute-2* results in arrested meristemoids (self-renewing stomatal precursors), a few stomata were formed in *mute-2* expressing MpSETA (Fig. 1b) (2.67 ± 1.56 and 2.33 ± 1.72 per abaxial side of the cotyledon in lines #8-4 and #10-11, respectively [mean $\pm$ s.d.; $n = 12$]). Notably, hydathode pores (a modified form of stomatal pores) were often found in these lines. This might be due to the high expression activity of the AtMUTE promoter in the hydathode of cotyledons¹⁸. MpSETA also exhibited the potential to rescue *fama-1*. *A. thaliana fama* mutant displays caterpillar-like stomatal-lineage cells called “*fama* tumors,” where terminal symmetric division occurs more than once⁸. In *fama-1* expressing MpSETA, excess cell divisions in the stomatal lineage were suppressed, although no mature stomata were formed (Extended Data Fig. 2b,c). Neither stomata nor stomatal-lineage cells were found in *spch-3* expressing MpSETA (Extended Data Fig. 2a). These results, showing that MpSETA is partially functional in stomatal cell division and differentiation in *A. thaliana*, suggest that it can interact with AtICE1 and AtSCRM2. Therefore, we tested this ability in yeast two-hybrid assays and bimolecular fluorescent complementation (BiFC) assays. We found that MpSETA physically interacted with AtICE1 and AtSCRM2 (Extended Data Fig. 2d,e). Thus, our findings indicate that MpSETA from the astomatous liverwort is a bona fide Ia bHLH TF, although its amino acid sequence is divergent.

MpSETA is preferentially expressed in developing sporophyte.

To investigate the expression pattern of MpSETA in *M. polymorpha*, we reanalyzed the public RNA-seq dataset from several organs^{4,19–22} and found that MpSETA was preferentially expressed in the diploid sporophyte, whereas its expression level was low in the haploid gametophyte (Fig. 2a).

We examined in detail the expression pattern of MpSETA by characterizing the different stages of development in the *M. polymorpha* sporophyte (Fig. 2b–c); in the wild type, sporophytes are divided into three tissues: foot, seta, and sporangium (Fig. 2b). The foot plays a key role in nutrient transport between the gametophyte and sporophyte, while the seta, which comprises files of elongated cells that form a stalk suspending the sporangium, plays a pivotal role in spore dispersal^{23–25}. After late spore maturation in sporophyte development, the seta extends through the elongation of its cells and thrusts the sporangium outwards, causing it to break through the surrounding calyptra and pseudoperianth, which are tissues derived from the archegonium and archegoniophore, respectively, to protect the sporophyte during development^{23–25}. As a result, spores are dispersed after sporangium dehiscence by desiccation. Based on the unique developmental events as shown in the previous studies^{22,23,26}, we divided the development of the sporophyte into 10 stages, as follows (Fig. 2c,d): (I) 2-cell stage with the epibasal cell (the upper cell that forms the foot and seta) and the hypobasal cell (the lower cell that forms the sporangium); (II) 4- or 8-cell stage; (III) early-globular stage with a distinct amphithecium (the outermost tissue that forms the capsule wall) and the endothecium (the

inner archesporial tissue); (IV) late-globular stage; (V) archesporial-tissue stage with the visible differentiated foot, seta, and archesporial tissues; (VI) sporogenous-cell stage with elaterocytes and sporogenous cells, which are the precursors of elaters and sporocytes (spore mother cells); (VII) sporocyte stage; (VIII) spore-tetrad stage (after meiosis); (IX) matured stage; (X) seta-elongated stage. Note that the seta and foot are established between stages V and IX. In stage VIII, we anatomically observed proliferative (symmetric) cell divisions in the seta region (Fig. 2e). Moreover, cell division activity was detected not in the foot but rather in the seta of sporophytes between stages VII and VIII using a G2-M phase reporter line, *proMpCYCB;1:Dbox-GUS*²⁷ (Fig. 2f). Thus, we concluded that the cell files in the seta are established by a few proliferative cell divisions of the putative “seta mother cell” in the later developmental stage.

We detected the promoter activity of *MpSETA* in the sporophyte by generating transformants expressing the β -glucuronidase gene (*GUS*) under the control of the *MpSETA* promoter (*proMpSETA:GUS*). *GUS* activity was found in developing sporophytes between stages IV and VII, especially in seta, whereas no *GUS* activity was found in stages VIII and IX (Fig. 2g). In gametophytic tissues, *GUS* signals were detected only in young antheridia (Extended Data Fig. 3). These findings suggest that *MpSETA* is expressed early in seta development and may regulate seta cell division and/or differentiation rather than setal cell elongation.

***Mpseta*^{ko} mutants show defects in setal formation.**

We generated loss-of-function mutants of *MpSETA* by homologous recombination-mediated gene targeting to reveal the function of *MpSETA* *in vivo*²⁸. Therefore, we obtained two independent *MpSETA* knock-out lines (*Mpseta-1*^{ko} and *Mpseta-2*^{ko}) in which the genomic regions encoding the bHLH domain were replaced with a hygromycin-resistance gene cassette (Extended Data Fig. 4a-c). We confirmed the loss of the full-length transcripts of the *MpSETA* gene by reverse-transcription polymerase chain reaction (RT-PCR) analysis of homozygous mutant sporophytes produced from crosses between *Mpseta-1*^{ko} or *Mpseta-2*^{ko} males and females (Extended Data Fig. 4d). Although *MpSETA* reporter gene expression was found in the developing antheridia (Extended Data Fig. 3), no obvious phenotype was observed during sperm formation in *Mpseta*^{ko} lines (Extended Data Fig. 4e), and mutant males were fertile.

We crossed males and females of *Mpseta*^{ko} and compared the resulting sporophytes with those of the wild type to investigate the phenotypes of *Mpseta*^{ko} mutants in the diploid generation. In longitudinal sections of the mature sporophytes of *Mpseta*^{ko}, we found anatomical defects in setal cell development (Fig. 3a,b). We did not observe any elongated setal cells or cell files of setal cells in *Mpseta*^{ko} mutants. The SF/SP ratio (a ratio of the length from the foot to the proximal side of sporangium to the total sporophyte length) and the number of cells around the seta region were significantly reduced in *Mpseta*^{ko} compared with the wild type (Fig. 3c,d). A detailed analysis of earlier stages of sporophyte development revealed that defects in *Mpseta-1*^{ko} setae could be found even at stage VI, the earliest stage at which putative seta mother cells are unequivocally recognized (Fig. 3e). Despite the obvious loss of setae, spores and other sporophytic tissues were normally formed in *Mpseta-1*^{ko} (Fig. 3e,f). We concluded that *Mpseta*^{ko} mutants have defects in the differentiation from putative seta precursor cells to seta mother cells, which prevents the induction of subsequent proliferative cell divisions. At one-month postfertilization, the wild-type sporangia were pushed out of the calyptras.

192 In contrast, *Mpseta*^{ko} sporangia remained buried inside the calyptras, presumably due to
 193 the defects in seta development, and hence were not exposed to the outside (Fig. 3g).
 194 Because sporangia are completely wrapped by calyptras and pseudoperianths, spore
 195 dispersal does not occur unless the sporangia are pushed out of the calyptras by seta cell
 196 elongation²⁹. In genetic complementation experiments, we generated a transgenic line
 197 with the genomic region of *MpSETA* introduced into *Mpseta-1*^{ko} (*gMpSETA Mpseta-1*^{ko}).
 198 The sporophytes obtained by crossing these complemented lines had normal setae, and
 199 their sporangia were pushed out of the calyptra as in the wild type (Fig. 3). Consequently,
 200 these results suggest that *MpSETA* is essential for the setal formation and spore dispersal
 201 in *M. polymorpha*.

202

203 **MpICE2 interacts with MpSETA and regulates seta development.**

204 Because the interaction between Ia and IIIb bHLHs is evolutionarily conserved^{9,10,30}, we
 205 examined whether Ia bHLH and IIIb bHLH physically interact with each other in *M.*
 206 *polymorpha*. We performed phylogenetic analyses to find IIIb bHLH genes in *M.*
 207 *polymorpha* and identified two genes encoding IIIb bHLH proteins in the genome of *M.*
 208 *polymorpha*, both of which are orthologous to *AtICE1* and *AtSCRM2* (Extended Data Fig.
 209 5). We named these genes *MpICE1* (Mp4g04910) and *MpICE2* (Mp4g04920). The amino
 210 acid sequences of the bHLH and ACT-like domains of IIIb bHLHs were highly conserved
 211 among land plants (Extended Data Fig. 6). Additionally, two putative IIIb bHLH-
 212 encoding genes, *LcICE1* and *LcICE2*, were found in the genome of *L. cruciata* (Extended
 213 Data Figs. 5 and 6).

214 Reanalyzing public RNA-seq data revealed high expression of *MpICE2* at 13
 215 days postfertilization in young sporophytes (from stage III to stage IV)¹⁹, in which
 216 putative seta mother cells were dividing and differentiating, whereas the expression level
 217 of *MpICE1* was almost constant in all tissues (Fig. 4a). Thus, we assumed that *MpICE2*
 218 may function predominantly in cooperation with *MpSETA* during seta cell formation
 219 rather than *MpICE1*. We expressed *Citrine* (a yellow fluorescent protein; YFP), *GUS*, and
 220 the nuclear localization-signal (*NLS*) fusion gene (*Citrine-GUS-NLS*) under the control
 221 of the *MpICE2* promoter to confirm the tissue/cell-level expression pattern of *MpICE2*
 222 in the sporophyte. In this line, GUS signals were detected overall in stages IV and V
 223 sporophytes, specifically in the seta and foot in stages VI–VIII sporophytes, and only in
 224 the foot in mature sporophytes together with Citrine and GUS signals in gametophytic
 225 tissues (Extended Data Fig. 7).

226 We performed yeast two-hybrid assays using full-length *MpSETA* in pairwise
 227 combinations with *MpICE2* to test the interaction between *MpSETA* and *MpICE2*; in this
 228 context, a physical interaction was observed (Fig. 4b). Additionally, a BiFC assay was
 229 used to test the interaction between these bHLH TFs. YFP signals were detected in the
 230 nuclei of *Nicotiana benthamiana* leaves coexpressing *MpSETA*-nYFP and *MpICE2*-
 231 cYFP (Fig. 4c). These results suggest that *MpSETA* can interact with *MpICE2* in *M.*
 232 *polymorpha*.

233 We generated two independent genome-edited lines using the CRISPR/Cas9
 234 system (*Mpice2-2*^{ge} and *Mpice2-6*^{ge}) to assess the function of *MpICE2*^{31,32}. The first one
 235 retained two amino acid substitutions (L503H and M504L) that were predicted to be
 236 important for the DNA-binding activity of the bHLH domain³³, whereas the second had
 237 a frame-shift mutation that caused deletion of the C-terminal half of the bHLH domain
 238 and the whole ACT-like domain (Extended Data Fig. 8). In these *Mpice2* mutants, the

setae of sporophytes were not formed (Fig. 4d,e), and the number of cells between the foot and sporangium was significantly reduced (Fig. 4f,g). Additionally, we found that the number of cells in the seta region of *Mpice2-2^{se}* exceeded that of *Mpice2-6^{se}* (Fig. 4g). This is probably because the predicted translational product of *Mpice2-2^{se}* has a two-amino acid substitution in the bHLH domain and might be partially functional in comparison with the null allele *Mpice2-6^{se}*. In both *Mpice2^{se}* mutants, the sporophytes did not emerge outside of the protective organs derived from the archegonia, similar to the *Mpseta^{ko}* mutants (Fig. 4h). Since the single mutants of *Mpice2^{se}* showed almost the same phenotype as that of the *Mpseta^{ko}* mutants, *MpICE1* may not be functionally redundant with *MpICE2*, at least in setal formation. The *Mpice2^{se}* phenotype in the setal region was completely suppressed by introducing the genomic region of *MpICE2* into *Mpice2-6^{se}* (Fig. 4d-h). Therefore, we can conclude that the MpSETA-MpICE2 heterodimer plays an important role in the setal development of *M. polymorpha*.

Next, we tested if *MpICE2* can enhance the rescue of the stomatal phenotype of *mute-2* by *proAtMUTE:MpSETA*. The overexpression of *MpICE2* in *proAtMUTE:MpSETA mute-2* did not enhance stomatal formation (Extended Data Fig. 9a) (3.63 ± 3.16 and 2.64 ± 1.41 per abaxial side of the cotyledon in the lines #8-4 and #10-11, respectively [mean $\pm$ s.d.; $n = 11$]). Therefore, the MpSETA-MpICE2 heterodimer does not appear to regulate the expression of stomatal genes in *A. thaliana*. Additionally, the expression of *MpICE1* or *MpICE2* under the control of the *AtICE1* promoter in *ice1-2 scrm2-1* mutants failed to cause stomatal-lineage cell formation (Extended Data Fig. 9b). Therefore, *MpICE1* and *MpICE2* cannot act with *AtSPCH* to regulate stomatal formation in *A. thaliana*, despite the similarity of their amino acid sequences with *AtICE1* (Extended Data Fig. 6).

Discussion

In this article, we showed that two transcription factors, MpSETA (Ia bHLH) and MpICE2 (IIIb bHLH), play a pivotal role in controlling the formation of the diploid tissue seta in the sporophyte of *M. polymorpha*, which is an astomatous liverwort (Figs. 3 and 4). Similarly, in other non-liverwort land plants, a module formed by subfamilies Ia and IIIb bHLH TFs regulates GRNs in stomatal development². *MpSETA* could partially complement the defects of *A. thaliana mute* and *fama*, suggesting similar properties of Ia bHLH TFs from *M. polymorpha* and *A. thaliana*. However, *MpSETA* was unable to complement the *spch* mutant (Fig. 1b and Extended Data Fig. 2). These results corroborate the previous hypothesis, stating that the ancestral Ia bHLH proteins had a MUTE- and FAMA-like function^{34,35}. Although the nature of *MpSETA*-expressing cells remains unknown, *MpSETA* may function as a regulator of cell differentiation and asymmetric cell division during the setal formation, similar to its role in stomatal formation (Fig. 5a).

Although the stomata and setae completely differ morphologically, both play a common role in promoting sporangial dehiscence and spore dispersal. In mosses and hornworts, stomata are present in the epidermis of the sporangia and function in the desiccation of the sporangia by gas exchange to promote its dehiscence and spore dispersal^{25,36}. In *A. thaliana*, *AtICE1* controls stomatal development on the anther epidermis and can regulate dehydration and dehiscence of the anther³⁷. However, the seta is the tissue that supports the sporangia and plays a role through cell elongation in thrusting the sporangia outside the surrounding maternal tissues to permit the long-distance dispersal of spores^{24,25}. In this study, we found that the lack of setae in *Mpseta^{ko}*

and *Mpice2^{ge}* mutants prevented sporangia dehiscence and spore dispersal due to the inability of the sporangia to break through the protective organs around it (Figs. 3 and 4). Thus, the Ia-IIIb bHLH module plays a common role in the development of the tissues involved in spore or pollen dispersal.

Previous research suggested that land plant evolution occurred through the reutilization and/or modification of preexisting GRNs^{20,38–43}. Here, we hypothesize that the Ia-IIIb bHLH module was primarily used in stomatal formation and secondarily co-opted to setal formation in the common ancestor of “Setaphyta” (Fig. 5b). Because only mosses and liverworts have setae, a moss–liverwort clade is called Setaphyta^{1,44,45}. However, the process of setal development differs between mosses and liverworts; whereas a transient intercalary meristem, called the seta meristem, produces seta in mosses, the body plan, including seta, foot, and sporangium, is established by the formative cell division at an early stage in liverworts²⁵. Therefore, whether the Ia-IIIb bHLH module is involved in the formation of seta in mosses, such as *P. patens*^{46,47}, should be tested.

In *A. thaliana*, the AtFAMA-AtICE1 heterodimer regulates not only the development of stomata but also of myrosin idioblasts, which are adjacent to vascular tissues and contribute to the defense against herbivores^{41,48,49}. Because myrosin cells are present only in the Brassicales, the AtFAMA-AtICE1 module for stomatal formation was co-opted for myrosin cell development during the evolution of Brassicales (Fig. 5b). The AtFAMA-AtICE1 module is thought to regulate the expression of different genes in stomatal and myrosin cell lineages^{41,48,49}; however, the detailed mechanism remains unknown. Liverworts have already lost several stomatal-related genes^{4,50}, such as the leucine-rich-repeat receptor-like gene *TOO MANY MOUTHS (TMM)* and the secreted peptide gene *EPIDERMAL PATTERNING FACTOR 1/2 (EPF1/2)*. The mechanisms underlying the development of setal cell lineage may differ from those related to stomatal cell lineages. AtMUTE directly regulates cell cycle-related genes (cyclin and cyclin-dependent kinase genes) and several stomatal-related TF genes⁵¹, and AtSPCH directly regulates stomatal-related genes and brassinosteroid pathway genes⁵². In *M. polymorpha*, orthologues of many of the Ia-IIIb bHLH target genes, including *CYCB*, *CYCD*, *CDKB*, *ERECTA*, and BZR/BES family TF genes, are conserved⁴. RNA-seq and ChIP-seq analyses could identify genes that function downstream of the MpSETA-MpICE2 module and help to clarify whether other stomatal formation-related genes are involved in the setal formation.

Are *SMF* genes and/or MpSETA orthologs conserved in the Jungermanniopsida or Haplomitriopsida liverworts? A BLAST search of 1,000 plant transcriptomes (OneKP)⁵³, using AtFAMA as a query, revealed the absence of Ia bHLH TFs in liverworts except for MpSETA. Because the liverwort transcriptome samples used in OneKP often do not contain sporophytes, detecting genes specifically or transiently expressed in developing sporophytes, such as Ia bHLH, is difficult. Additionally, owing to the loss of seta in the Ricciaceae species, investigating whether these species have Ia bHLH and whether the Ia bHLHs are functional to understand seta evolution is crucial. Genome analyses of various plant species in the future could be useful for understanding the relationship between Ia bHLH diversification and stomata/seta formation.

Methods

Phylogenetic analysis

333 The bHLHs were classified according to Pires and Dolan⁵. We retrieved the amino acid
 334 sequence information from each of the following; MarpolBase (<https://marchantia.info/>),
 335 Phytozome v.13 (<https://phytozome-next.jgi.doe.gov/>), OneKP
 336 (<https://db.cngb.org/onekp/>), TAIR (<http://www.arabidopsis.org/>), and NCBI
 337 (<https://www.ncbi.nlm.nih.gov/genome/?term=PRJNA701193>). Then, we aligned the
 338 bHLH domain and the C-terminal ACT-like domain using MAFFT⁵⁴ v.6.864
 339 (<https://www.genome.jp/tools-bin/mafft>) with the default parameters. Noting that the
 340 alignment gaps were removed manually. Next, we visualized the amino acid sequences
 341 with Jalview⁵⁵ v.2.11.2.1. To predict the importance of the amino acids in nucleotide
 342 binding and dimer formation in the bHLH domains, we used BLAST searches of NCBI
 343 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Alternately, we performed the phylogenetic tree
 344 constructions using the maximum-likelihood algorithm on MEGA 7⁵⁶ with the JTT+G+I
 345 and LG+G+I substitution models for Ia and IIIb bHLH, respectively. To assess the
 346 statistical support for the topology, we performed bootstrap analyses with 1,000 replicates
 347 for each analysis. Subfamilies III(a+c) and III(d+e) bHLHs were chosen as the outgroup
 348 of the phylogeny for Ia and IIIb bHLH, respectively.

349 A BLAST search of OneKP (<https://db.cngb.org/onekp/>), using AtFAMA as a
 350 query, revealed the absence of Ia bHLH TFs in liverworts except for MpSETA.

351

352 **Plant materials and growth conditions**

353 We asexually maintained the male and female accessions of *M. polymorpha* L.,
 354 Takaragaike-1 (Tak-1), and Tak-2, respectively. A female progeny backcrossed to Tak-1
 355 for three backcross generations (BC3-38) was also used as the wild type. We cultured
 356 gemmae and thalli on half-strength Gamborg's B5 media containing 1% (w/v) agar and
 357 1% (w/v) sucrose under 50–60 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ continuous white light at 22°C. To
 358 induce gametangiophore development, we transferred gemmae to 16-h-light/8-h-dark
 359 conditions with 50–60 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ white light and 50–60 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$
 360 far-red light emitted from diodes (IR LED STICK 18W, Namoto) at 18°C and incubated
 361 for one month.

362 We used the *Arabidopsis thaliana* Columbia-0 (Col-0) accession as wild type
 363 except for *mute-2*, where Wassilewskija-4 (Ws-4) was used. The seeds were surface
 364 sterilized with 70% ethanol, then sown onto half-strength MS media containing 0.5%
 365 (w/v) gellan gum and 1% (w/v) sucrose. We incubated the seeds at 22°C under 50–60
 366 $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ continuous white light.

367

368 **Complementation tests of *A. thaliana* stomatal defective mutants**

369 We obtained T-DNA insertion mutants *spch-3* (SAIL_36_B06) and *fama-1*
 370 (SALK_100073) from the Arabidopsis Biological Resource Center (ABRC), while *mute-2*
 371 (FLAG_225D03) was from the French National Institute for Agricultural Research
 372 (INRA). *ice1-2 scrm2-1*⁹ was provided by K.U. Torii. To construct *proAtSPCH:MpSETA*,
 373 firstly, we amplified the AtSPCH promoter (2,572 bp upstream of the translational
 374 initiation site) and MpSETA CDS from Col-0 gDNA and cDNA derived from Tak-1
 375 antheridiophores, respectively. Then, we fused *proAtSPCH* and MpSETA CDS fragments
 376 by PCR using PrimeSTAR GXL polymerase (Takara Bio). The resultant PCR fragment
 377 was subcloned into pENTR1A entry clones (Invitrogen) at the Sall and EcoRV restriction
 378 sites using an In-Fusion HD Cloning Kit (Takara Bio), which we transferred into the
 379 destination vector pGWB501⁵⁷ using Gateway LR Clonase II Enzyme mix (Thermo

Fisher Scientific). We generated the *proAtMUTE:MpSETA* and *proAtFAMA:MpSETA* constructs by LR recombination of the R4pGWB501⁵⁸ or R4pGWB601⁵⁸ with a pENTR1A containing the *MpSETA* CDS at the EcoRI sites and either pENTR5'/TOPO_ *proAtMUTE* harboring *proAtMUTE*, 1,930 bp upstream of the translational initiation site, or pENTR5'/TOPO_ *proAtFAMA* harboring *proAtFAMA*, 3,105 bp upstream of the translational initiation site (R4pGWB501_ *proAtMUTE:MpSETA* and R4pGWB601_ *proAtFAMA:MpSETA*). To construct *proAtICE1:MpICE1* or *proAtICE1:MpICE2*, we amplified, respectively, the *AtICE1* promoter (2,578 bp upstream of the translational initiation site) from Col-0 gDNA, and *MpICE1* or *MpICE2* CDS from cDNA derived from Tak-1 thalli. We fused *proAtICE1* and *MpICE1* or *MpICE2* CDS fragments by PCR using PrimeSTAR GXL polymerase. We subcloned the resultant PCR fragment into pENTR1A entry clones at the SalI and EcoRI restriction sites using an In-Fusion HD Cloning Kit and transferred them into the destination vector pGWB501⁵⁷ using Gateway LR Clonase II Enzyme mix. We introduced the resultant plasmids into *spch-3/+*, *mute-2/+*, *fama-1/+*, or *ice1-2/+ scrm2-1* heterozygous plants by the previously described method⁵⁹ using *Agrobacterium tumefaciens* strain GV3101. We confirmed that all the transformants had a single insertion event by segregation analyses. We used T₃ or T₄ homozygous plants. For *MpICE2* overexpression analyses, we transferred the *MpICE2* CDS into the destination vector pFAST-R02⁶⁰ using Gateway Clonase II Enzyme mix, and the resultant plasmid was introduced into *Ws-4*, *mute-2/+*, and *proAtMUTE:MpSETA mute-2/+* (#8-4 and #10-11). T₁ seeds expressing TagRFP were selected, and T₁ plants were used for the analyses. We stained the cotyledons with FM4-64 and observed them using an LSM780 laser scanning microscope (Carl Zeiss). Images were processed with Fiji v.1.0 (NIH). The sequences of the primers used in this study are shown in Supplementary Tables 1 and 2.

Gene expression analysis

Publicly available transcriptome data were downloaded from the Sequence Read Archive (SRA) repository. Accession numbers included; sporelings (SRR4450254, SRR4450255, SRR4450256)⁴, male thalli (DRR118949, DRR118950, DRR118951)²⁰, female thalli (DRR118943, DRR118944, DRR118945)²⁰, antheridiophores (DRR050346, DRR050347, DRR050348)²¹, archegoniophores (DRR050351, DRR050352, DRR050353)²¹, antheridia (DRR050349, DRR050350)²¹, archegonia (DRR209029, DRR209030, DRR209031, DRR209031)²², 13 DPF sporophytes (SRR1553297, SRR1553298, SRR1553299)¹⁹, and mature sporophytes (SRR896223)⁴. We pre-processed the RNA-seq data to filter out low-quality sequences using fastp⁶¹ v.0.20.0. The sequence reads were mapped to the *M. polymorpha* genome v.6.1 (<https://marchantia.info>) by STAR⁶² v.2.7.8a with default parameters. We performed the post-processing of SAM/BAM files using SAMtools⁶³ v.1.11. To calculate the transcript per million (TPM) via the RSEM⁶⁴ v.1.3.0 with default parameters, we used the read for each gene. We created the plots using Microsoft Office Excel for Mac (Microsoft) and Rstudio v.1.4.1106 (<https://www.rstudio.com/>).

Histology

We stained the sporophytes (stage I and II) with 4',6-diamidino-2-phenylindol (DAPI) as described previously²² and observed them using an LSM780 laser scanning microscope. We fixed the plant samples with 2% (w/v) paraformaldehyde and 2% (v/v)

glutaraldehyde in 0.05 M cacodylate buffer (pH 7.4) for 2 h at room temperature, post-fixed with 2% (v/v) osmium tetroxide in 0.1 M cacodylate buffer for another 2 h at room temperature, dehydrated in an ethanol series, substituted with acetone, and then embedded in Spurr's resin (Polysciences). We cut the Spurr's blocks into semi-thin sections (0.75–2 μ m) with glass knives on an ultramicrotome Leica Ultracut UCT (Leica Microsystems) stained with a solution containing 1% (w/v) sodium tetra-borate and 1% (w/v) toluidine blue O. The sections were mounted on MAS-coated glass slides (Matsunami Glass). In the end, we obtained the images using a VB-7010 (KEYENCE)/AxioCam HRc (Zeiss) and processed them with either Fiji or Adobe Photoshop Elements 9 (Adobe Systems).

Histochemical GUS staining

To construct *proMpSETA:GUS*, we amplified the genomic fragment of the 4,194 bp upstream region of the translational initiation site from Tak-1 gDNA using PrimeSTAR Max DNA polymerase (Takara Bio), subcloned it into pENTR1A at the EcoRI restriction sites using the In-Fusion HD Cloning Kit and then transferred it into the destination vector pMpGWB104⁶⁵ using Gateway LR Clonase II Enzyme mix (Thermo Fisher Scientific). To construct *proMpICE2:Citrine-GUS-NLS*, firstly we amplified the genomic fragment containing the 3,060 bp upstream region of the translational initiation site from Tak-1 gDNA, *Citrine* ORF from pMpGWB107⁶⁵, and *GUS-NLS* ORF from pPZP211_35S-NG-GUS-NLS-nosT⁶⁶. These fragments were fused by PCR, subcloned into pENTR1A at the SalI and EcoRV sites using In-Fusion HD Cloning Kit, and transferred into pMpGWB101⁶⁵. We introduced the resultant plasmids into Tak-1 accession by the previously described method⁶⁷ using *A. tumefaciens* strain GV2260.

The tissues of *proMpSETA:GUS* or *proMpICE2:Citrine-GUS-NLS* plants were vacuum-infiltrated and incubated at 37°C overnight in GUS staining solution containing 10 mM EDTA (pH 8.0), 100 mM NaH₂PO₄ (pH 7.0), 0.1% (v/v) Triton X-100, 0.5 g L⁻¹ 5-bromo-4-chloro-3-indolyl- β -D-glucuronic acid (X-Gluc), 5 mM potassium-ferrocyanide, and 5 mM potassium-ferricyanide. We washed the samples in 70% (v/v) ethanol and cleared them with chloral-hydrate / glycerol solution.

Generation of *Mpseta*^{ko} mutants

We amplified the Tak-1 genomic sequences 3,125 bp upstream and 3,101 bp downstream of the *MpSETA* bHLH domain coding region by PCR with a PrimeSTAR Max DNA polymerase and inserted them into the PacI and AscI sites of pJHY-TMp1²⁸, respectively. We introduced the vector into germinating F₁ spores from Tak-1 and Tak-2 crosses via *A. tumefaciens* strain GV2260 as previously described⁶⁸. We selected the transformed T₁ plants carrying the targeted insertion by PCR using GoTaq DNA polymerase (Promega). As T₁ plants of *Mpseta-1*^{ko} and *Mpseta-2*^{ko} were both females, we obtained the male mutants from F₁ sporelings by crossing female mutants with Tak-1.

Reverse transcription PCR

For gene expression analysis, we collected 21 DPF sporophytes from wild-type, *Mpseta-1*^{ko}, and *Mpseta-2*^{ko}. We extracted the total RNA by RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's protocol. We evaluated total RNA qualitatively and quantitatively using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific). First-strand cDNA was synthesized using ReverTra Ace qPCR RT Master Mix with

473 gDNA Remover (Toyobo), and semiquantitative RT-PCR was undertaken using Mp*EF1α*
474 as a loading control⁶⁹.

475

476 **Yeast two-hybrid assay**

477 We amplified the coding sequences of Mp*SETA* and Mp*ICE2* from cDNA derived from
478 mRNA of Tak-1 thalli by PCR using PrimeSTAR Max DNA polymerase or PrimeSTAR
479 GXL polymerase (Takara Bio). Also, we amplified the coding sequences of At*ICE1* and
480 At*SCRM2* from cDNA derived from Col-0 leaves using PrimeSTAR GXL polymerase.
481 We subcloned the resultant PCR fragments into pENTR1A at the EcoRI sites or Sall and
482 EcoRV sites using an In-Fusion HD Cloning Kit. To generate a bait destination vector,
483 pDEST-GBKT7-Amp^r, *Amp^r* was amplified from pDEST-GADT7⁷⁰ by PCR using
484 PrimeSTAR Max DNA polymerase, and the fragment was cloned into the SfoI site of
485 pDEST-GBKT7⁷⁰. The inserted fragments, Mp*SETA*, Mp*ICE2*, At*ICE1*, and At*SCRM2*,
486 were transferred into pDEST-GADT7 and/or pDEST-GBKT7-Amp^r using Gateway LR
487 Clonase II Enzyme mix. Bait and prey constructs were co-transformed into the yeast
488 strain Y2HGold (Clontech) using the Frozen-EZ Yeast Transformation II Kit (Zymo
489 Research) and the transformants were grown on the solid SD media lacking Leu and Trp
490 (SD-LW). To examine the interaction between the bait and prey proteins, transformants
491 were grown on the solid SD media lacking Leu, Trp, His, and adenine with 40 mg L⁻¹ X-
492 α-gal and 200 μg L⁻¹ Aureobasidin A (SD-LWHA/X/AbA) at 30°C. pDEST-GBKT7-
493 Amp^r and pDEST-GADT7 were used as negative controls (empty).

494

495 **BiFC**

496 The coding sequences of Mp*SETA*, Mp*ICE2*, At*ICE1*, and At*SCRM2* subcloned into
497 pENTR1A, as described above, were transferred into the pB4GWnY and/or
498 pB4GWcY/pB4cYGW binary vector⁷¹ using LR reaction to be fused with N-terminal or
499 C-terminal half of YFP (Mp*SETA*-nYFP, Mp*ICE2*-cYFP, cYFP-At*ICE1*, and cYFP-
500 At*SCRM2* driven by the *CaMV35S* promoter). Transformed *A. tumefaciens* strain
501 GV3101 cells harboring expression vectors were cultured and resuspended in distilled
502 water to a final optical density of OD₆₀₀ = 1.0. Mixed *Agrobacterium* cultures were
503 infiltrated into 4-week-old *Nicotiana benthamiana* leaves. The nuclei were stained with
504 DAPI at the 1 mg L⁻¹ concentration for 30 min. We observed the samples 1.5 days post-
505 inoculation (DPI) by an LSM780 laser scanning microscope. pB4GWnY and
506 pB4GWcY/pB4cYGW were used as negative controls (empty).

507

508 **Generation of Mpice2^{se} mutants**

509 To generate Mpice2^{se} mutants, we edited the Mp*ICE2* locus encoding the bHLH domain
510 using a CRISPR/Cas9 based genome-editing system, as previously described³¹. Two
511 sgRNAs were designed to generate Mpice2^{se} mutants. The oligonucleotides encoding
512 sgRNA were cloned into pMpGE_En03³¹ between BsaI sites, then introduced into
513 pMpGE010³¹. Male and female mutants that did not harbor T-DNA containing *Cas9* were
514 obtained from F₁ sporelings crossed with wild type and Mpice2-2^{se} or Mpice2-6^{se}.

515

516 **Complementation tests of Mpseta^{ko} and Mpice2^{se}**

517 To construct gMp*SETA* for Mpseta-1^{ko} complementation, the genomic region containing
518 the 4,194 bp upstream region and coding sequences was amplified from Tak-1 genomic
519 DNA. The fragment was cloned into pENTR1A between the EcoRI sites and then

introduced into pMpGWB301⁶⁵. The resultant plasmids were introduced into female *Mpseta-1^{ko}*. Male gMpSETA *Mpseta-1^{ko}* was obtained from F₁ sporelings produced from crosses between female gMpSETA *Mpseta-1^{ko}* and male *Mpseta-1^{ko}*. To construct gMpICE2 for *Mpice2-6^{ge}*, the genomic region containing 3,060 bp upstream regions and coding sequences was amplified from Tak-1 genomic DNA. The fragment was cloned into pENTR1A between the SalI and EcoRV sites and then introduced into pMpGWB101⁶⁵. The resultant plasmids were introduced into female *Mpice2-6^{ge}*. A male gMpICE2 *Mpice2-6^{ge}* was obtained from an F₁ sporeling derived from a cross between a female gMpICE2 *Mpice2-6^{ge}* and a male *Mpice2-6^{ge}*.

Data Availability

Public RNA-seq data can be downloaded from SRA repository (<https://www.ncbi.nlm.nih.gov/sra>) under accession PRJNA350270, PRJDB6579, PRJDB4420, PRJDB9329, and PRJNA265205. All data supporting the conclusions in the paper are available within the paper or the supplementary materials. Source data are provided with this paper. Sequence data and gene ID can be found in the database, as follows: MarpolBase (<https://marchantia.info>), Phytozome v.13 (<https://phytozome-next.jgi.doe.gov/>), OneKP (<https://db.cngb.org/onekp/>), TAIR (<http://www.arabidopsis.org/>), and NCBI (<https://www.ncbi.nlm.nih.gov/genome/?term=PRJNA701193>). Source data are provided with this paper.

Acknowledgements

We thank Tsuyoshi Nakagawa (Shimane University, Japan), Shoji Mano (National Institute for Basic Biology, Japan), Shigeo S. Sugano (National Institute of Advanced Industrial Science and Technology, Japan), and Keiko U. Torii (The University of Texas at Austin, USA) for sharing the materials. We also thank Keiji Nakajima (Nara Institute of Science and Technology, Japan) for sharing the figures of plants in Fig. 5b. We are grateful to James Raymond for his critical readings of this manuscript. This work was supported by MEXT/JSPS KAKENHI grants to M.S. (JP19K06722 and JP20H05416), to K.T. (JP26711017 and JP18K06283), to Y.O. (JP18K19964), to T.M. (JP20H05905 and JP20H05906), to I.H.-N. (JP15H05776), to R.N. (JP20H04884) and T.S. (JP18K06284); Grants-in-Aid JSPS Fellows to K.C.M. (JP21J14990) and M.S. (JP12J05453) and; the Takeda Science Foundation, the Kato Memorial Bioscience Foundation, and the Ohsumi Frontier Science Foundation to M.S. J.L.-M. and Y.-T.L. were supported by Ph.D. studentships from the Darwin Trust of Edinburgh.

Author Contributions Statement

K.C.M. and T.S. conceived and designed the research in general; K.C.M. performed most of the experiments and analyzed the data; M.S. and Y.M. performed the experiments on *MpSETA*; J.L.-M., Y.-T. L., G.I., and J.G. performed the experiments on *MpICE2*; K.T., Y.O., T.M., I.H.-N., R.N., J.G., and T.K. supervised the experiments; K.C.M. and T.S. wrote the manuscript; All authors read, edited, and approved the manuscript.

Competing Interests Statements

567 The authors declare no competing interests.

568

569 Figure legends

570 Fig. 1 | MpSETA is the only bHLH transcription factor that belongs to subfamily Ia

571 in *M. polymorpha*. **a**, A maximum-likelihood bHLH phylogenetic tree of subfamilies Ia,
572 Ib(1) (purple), II (gray), and III(a+c) (outgroup) is shown. Numbers at branches indicate
573 bootstrap values calculated from 1,000 replicates. Ia bHLHs are divided into four groups:
574 TEC1/6 clade (light blue), MYC70 clade (cream-yellow), SMF/FAMA clade (magenta),
575 and SPCH/MUTE clade (green). Species are abbreviated as follows: Mp, *M. polymorpha*
576 (liverwort); Lc, *L. cruciata* (liverwort); Pp, *P. patens* (moss); Cepur, *Ceratodon purpureus*
577 (moss); Agr, *Anthoceros agrestis* (hornwort); Sm, *Selaginella moellendorffii*
578 (lycophyte); AmTr, *Amborella trichopoda* (basal angiosperm); Os, *Oryza sativa*
579 (monocot); At, *A. thaliana* (dicot). An arrow indicates MpSETA/MpBHLH35
580 (Mp2g04160), and arrowheads indicate bHLH TFs mentioned as FAMA-like bHLH TFs
581 in a previous study¹⁴. Amino acid sequences from only *A. thaliana* and *M. polymorpha*
582 were used for subfamilies Ib(1), II, and III(a+c). **b**, Confocal images of *A. thaliana* abaxial
583 cotyledons of wild type (Ws-4), *mute-2*, and *proAtMUTE:MpSETA mute-2* at 9 days after
584 stratification (DAS). The upper and lower panels show the middle and tip areas of the
585 cotyledons, respectively (left image). Arrowheads and asterisks indicate stomata and
586 hydathode pores, respectively. The experiments were repeated independently three times
587 with similar results. Bars, 100 μ m (**b**, confocal images), and 1 mm (**b**, left).

588

589 Fig. 2 | MpSETA is preferentially expressed in developing sporophyte. **a**, Box plot

590 showing the expression profiles of MpSETA across nine *M. polymorpha* tissues. The Y-
591 axis shows transcripts per million (TPM). Exact sample sizes (*n*) are shown above the
592 boxes. **b**, Tissue section of mature *M. polymorpha* sporophyte. **c**, Developmental stages
593 of sporophytes after 4–40 days' postfertilization. (I) sporophyte differentiating an epibasal
594 cell and a hypobasal cell, (II) 4- or 8-cell sporophyte, (III) sporophyte differentiating
595 amphithecium and endothecium, (IV) later globular sporophyte, (V) sporophyte
596 differentiating foot, seta, and archesporial tissues, (VI) sporophyte differentiating
597 sporogenous cells and elaterocytes, (VII) sporophyte differentiating sporocytes, (VIII)
598 sporophyte differentiating spore tetrads, (IX) mature sporophyte before seta elongation,
599 (X) mature sporophyte after seta elongation. Arrows and arrowheads indicate the
600 endothecium and the cell wall of the first cell division, respectively. **d**, Magnified images
601 showing the foot and seta for each stage of sporophytes. The same images as used in (**c**).
602 **e**, The formation of the cell files of seta in the wild type. Enlarged images of the square
603 areas are shown in (**d**). Arrowheads indicate the plane of proliferative cell divisions. **f**,
604 Histochemical detection of β -glucuronidase (GUS) activity in the sporophytes crossed
605 with male wild type and female *proMpCYCB;1:Dbox-GUS*. **g**, Histochemical detection of
606 GUS activity driven by the MpSETA promoter in the developing seta region. f, foot; s,
607 seta; at, archesporial tissue; sp, sporangium; ca, calyptra; p, pseudoperianth. Arrowheads
608 indicate the cell wall of the first cell division. The experiments in this figure were repeated
609 independently at least three times with similar results. All the boxplots show the
610 distribution of all samples with dots, the median (center line) and the first and third
611 quartile (upper and lower hinges), and the whiskers extend to a maximum of 1.5 $\times$ the
612 interquartile range. Bars, 50 μ m (**c**, top), and 100 μ m (**b**, **c**, bottom, **d**, **f** and **g**).

613

Fig. 3 | *Mpseta*^{ko} mutants show developmental defects in the seta. **a,b**, Longitudinal sections of sporophytes of the indicated genotypes. Enlarged images of the square areas are shown in **(b)**. Dashed and solid arrows indicate the length of the sporophyte (SP) and the length from the boundary between the gametophyte and sporophyte to the proximal side of the sporangium (SF), respectively. **c,d**, Quantitative data showing the SF/SP length ratio and the number of cells in the seta region. Since the boundary between the seta and foot is unclear, cells excluding transfer cells (the cells at the boundary between gametophyte and sporophyte) were used to count the number of cells. The different letters indicate significantly different mean values at $p < 0.01$ (two-sided Tukey's HSD test, $n = 6$ independent lines). **e**, Comparative analyses of the tissue development between wild type (WT) and *Mpseta-1*^{ko}. Enlarged images of longitudinal sections of stage VI, stage VII, and stage VIII sporophytes. **f**, Sporogenesis process in wild type (WT) and *Mpseta-1*^{ko}. Sporocytes are the cells before the pre-meiotic stage, and spore tetrads are the cells immediately after meiosis. **g**, One month-postfertilization archegoniophores that produce mature sporophytes. Arrowheads indicate sporangia exposed to the outside of the calyptra due to the elongation of seta cells. The experiments in this figure were repeated independently at least three times with similar results. All the boxplots show the distribution of all samples with dots, the median (center line) and the first and third quartile (upper and lower hinges), and the whiskers extend to a maximum of 1.5× the interquartile range. Bars, 50 μ m (**f**), 100 μ m (**a**, **b**, **e**), and 5 mm (**e**).

Fig. 4 | *MpICE2* regulates setal development in cooperation with *MpSETA*. **a**, Box plot showing the expression profiles of *MpICE1* and *MpICE2* across nine *M. polymorpha* tissues. The Y-axis shows TPM. Exact sample sizes (n) are shown above the boxes. **b**, Y2H assays showing the interaction between *MpSETA* and *MpICE2*. *MpSETA* fused with GAL4 DBD was used as bait, whereas *MpICE2* fused with GAL4 AD was used as prey. DBD alone and AD alone were used as negative controls. **c**, BiFC assays showing the interaction between *MpSETA* and *MpICE2* in *N. benthamiana*. *MpSETA* was fused to nYFP, whereas *MpICE2* was fused to cYFP. nYFP alone and cYFP alone were used as the negative controls. Nuclei were stained by 4',6-diamidino-2-phenylindole (DAPI). **d,e**, Longitudinal sections of sporophytes of indicated genotypes. Enlarged images of the square areas are shown in **(e)**. Dashed and solid arrows indicate SP and SF, respectively. **f,g**, Quantitative data showing the SF/SP length ratio and the number of cells in the seta region. Because the boundary between the seta and foot is unclear, cells excluding transfer cells were used to count the number of cells. The different letters indicate significantly different mean values at $p < 0.01$ (two-sided Tukey's HSD test). $n = 6$ (*Mpice2-2*^{ge} and *Mpice2-6*^{ge}) or 7 (WT and *gMpICE2 Mpice2-6*^{ge}). **h**, One month-postfertilization archegoniophores that produce mature sporophytes. Arrowheads indicate sporangia exposed to the outside of the calyptra due to the elongation of seta cells. The experiments in this figure were repeated independently at least three times with similar results. All the boxplots show the distribution of all samples with dots, the median (center line) and the first and third quartile (upper and lower hinges), and the whiskers extend to a maximum of 1.5× the interquartile range. Bars, 10 μ m (**c**, YFP, DAPI, and Merge), 100 μ m (**c**, left, **d**, and **e**), and 5 mm (**f**).

Fig. 5 | Function and co-option of the Ia-IIIb bHLH module during land plant evolution. **a**, Schematic model comparing the molecular functions of the Ia-IIIb bHLH

TF modules in *A. thaliana* and *M. polymorpha* during cell fate determination. In *A. thaliana*, the heterodimers of Ia bHLHs (SPCH, MUTE, and FAMA) and IIIb bHLHs (ICE1 and SCRM2) control stomatal development. In *M. polymorpha*, the heterodimer of Ia bHLH (MpSETA) and IIIb bHLH (MpICE2) regulates cell differentiation and cell division in the seta precursor transition and might be directly or indirectly involved in the symmetric division of a putative seta mother cell. **b**, An evolutionary model for the Ia-IIIb bHLH TF module. The co-option of the Ia-IIIb bHLH module might have occurred multiple times independently during land plant evolution. First, stomata and a transcriptional module comprising Ia-IIIb bHLHs evolved in the common ancestor of land plants. In the ancestral plant, the Ia bHLHs may have had MUTE- and FAMA-like functions. Second, the Ia-IIIb bHLH TF module might have been co-opted to regulate setal development in the ancestor of the Setaphyta. Third, after mosses and liverworts diverged, the common ancestor of liverworts lost its stomata. A co-option of the Ia-IIIb bHLH module occurred in the Brassicales plants for regulating myrosin idioblast development.

References

1. Ligrone, R., Duckett, J. G. & Renzaglia, K. S. Major transitions in the evolution of early land plants: A bryological perspective. *Ann. Bot.* **109**, 851–871 (2012).
2. Romani, F. & Moreno, J. E. Molecular mechanisms involved in functional macroevolution of plant transcription factors. *New Phytol.* **230**, 1345–1353 (2021).
3. Catarino, B., Hetherington, A. J., Emms, D. M., Kelly, S. & Dolan, L. The stepwise increase in the number of transcription factor families in the Precambrian predated the diversification of plants on land. *Mol. Biol. Evol.* **33**, 2815–2819 (2016).
4. Bowman, J. L. *et al.* Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* **171**, 287–304.e15 (2017).
5. Pires, N. & Dolan, L. Origin and diversification of basic-helix-loop-helix proteins in plants. *Mol. Biol. Evol.* **27**, 862–874 (2010).
6. MacAlister, C. A., Ohashi-Ito, K. & Bergmann, D. C. Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. *Nature* **445**, 537–540 (2007).
7. Pillitteri, L. J., Sloan, D. B., Bogenschutz, N. L. & Torii, K. U. Termination of asymmetric cell division and differentiation of stomata. *Nature* **445**, 501–505 (2007).
8. Ohashi-Ito, K. & Bergmann, D. C. *Arabidopsis* FAMA controls the final proliferation/differentiation switch during stomatal development. *Plant Cell* **18**, 2493–2505 (2006).
9. Kanaoka, M. M. *et al.* SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell* **20**, 1775–1785 (2008).
10. Chater, C. C. *et al.* Origin and function of stomata in the moss *Physcomitrella patens*. *Nat. Plants* **2**, 16179 (2016).
11. Caine, R. S., Chater, C. C. C., Fleming, A. J. & Gray, J. E. Stomata and sporophytes of the model moss *Physcomitrium patens*. *Front. Plant Sci.* **11**, 643 (2020).
12. Chater, C. C. C., Caine, R. S., Fleming, A. J. & Gray, J. E. Origins and evolution of stomatal development. *Plant Physiol.* **174**, 624–638 (2017).

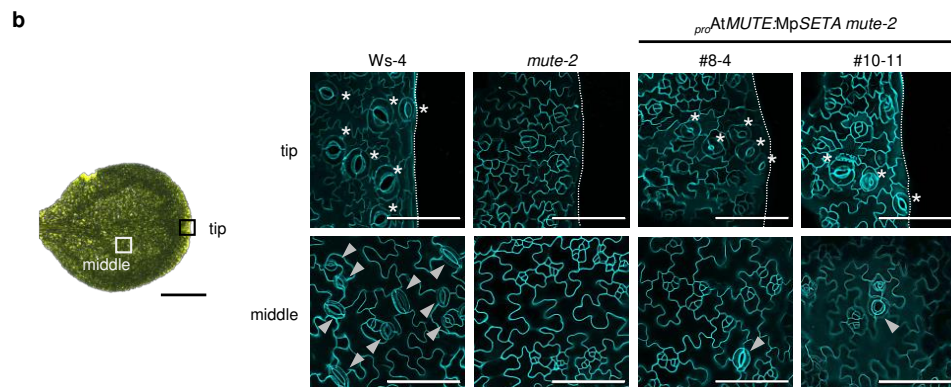
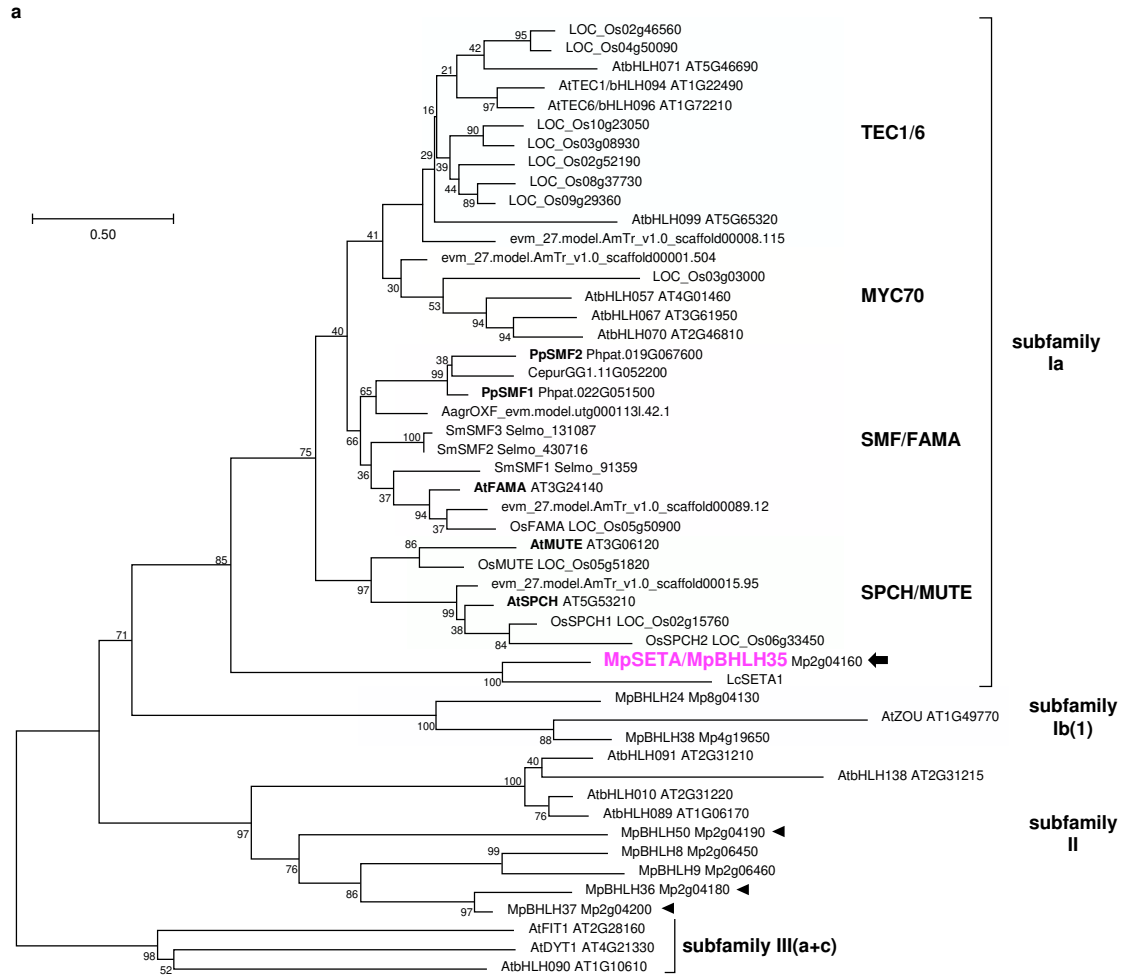
- 708 13. Harris, B. J., Harrison, C. J., Hetherington, A. M. & Williams, T. A. Phylogenomic
709 evidence for the monophyly of bryophytes and the reductive evolution of stomata.
710 *Curr. Biol.* **30**, 2001–2012.e2 (2020).
- 711 14. Flores-Sandoval, E., Romani, F. & Bowman, J. L. Co-expression and
712 transcriptome analysis of *Marchantia polymorpha* transcription factors supports
713 class C ARFs as independent actors of an ancient auxin regulatory module. *Front.*
714 *Plant Sci.* **9**, 1345 (2018).
- 715 15. Feller, A., Hernandez, J. M. & Grotewold, E. An ACT-like domain participates in
716 the dimerization of several plant basic-helix-loop-helix transcription factors. *J.*
717 *Biol. Chem.* **281**, 28964–28974 (2006).
- 718 16. Seo, H. *et al.* Intragenic suppressors unravel the role of the SCREAM ACT-like
719 domain for bHLH partner selectivity in stomatal development. *Proc. Natl. Acad.*
720 *Sci. U. S. A.* **119**, e2117774119 (2022).
- 721 17. Linde, A. M., Eklund, D. M., Cronberg, N., Bowman, J. L. & Lagercrantz, U. Rates
722 and patterns of molecular evolution in bryophyte genomes, with focus on complex
723 thalloid liverworts, Marchantiopsida. *Mol. Phylogenet. Evol.* **165**, 107295 (2021).
- 724 18. Pillitteri, L. J., Bogenschutz, N. L. & Torii, K. U. The bHLH protein, MUTE,
725 controls differentiation of stomata and the hydathode pore in Arabidopsis. *Plant*
726 *Cell Physiol.* **49**, 934–943 (2008).
- 727 19. Frank, M. H. & Scanlon, M. J. Transcriptomic evidence for the evolution of shoot
728 meristem function in sporophyte-dominant land plants through concerted selection
729 of ancestral gametophytic and sporophytic genetic programs. *Mol. Biol. Evol.* **32**,
730 355–367 (2015).
- 731 20. Yamaoka, S. *et al.* Generative cell specification requires transcription factors
732 evolutionarily conserved in land plants. *Curr. Biol.* **28**, 479–486.e5 (2018).
- 733 21. Higo, A. *et al.* Transcriptional framework of male gametogenesis in the liverwort
734 *Marchantia polymorpha* L. *Plant Cell Physiol.* **57**, 325–338 (2016).
- 735 22. Hisanaga, T. *et al.* Deep evolutionary origin of gamete-directed zygote activation
736 by KNOX/BELL transcription factors in green plants. *eLife* **10**, e57090 (2021).
- 737 23. Shimamura, M. *Marchantia polymorpha*: Taxonomy, phylogeny and morphology
738 of a model system. *Plant Cell Physiol.* **57**, 230–256 (2016).
- 739 24. Haig, D. Filial mistletoes: The functional morphology of moss sporophytes. *Ann.*
740 *Bot.* **111**, 337–345 (2013).
- 741 25. Ligrone, R., Duckett, J. G. & Renzaglia, K. S. The origin of the sporophyte shoot
742 in land plants: A bryological perspective. *Ann. Bot.* **110**, 935–941 (2012).
- 743 26. Durand, E. J. The development of the sexual organs and sporogonium of
744 *Marchantia polymorpha*. *Bull. Torrey Bot. Club* **35**, 321–335 (1908).
- 745 27. Hernández-García, J. *et al.* Coordination between growth and stress responses by
746 DELLA in the liverwort *Marchantia polymorpha*. *Curr. Biol.* **31**, 3678–3686.e11
747 (2021).
- 748 28. Ishizaki, K., Johzuka-Hisatomi, Y., Ishida, S., Iida, S. & Kohchi, T. Homologous
749 recombination-mediated gene targeting in the liverwort *Marchantia polymorpha*
750 L. *Sci. Rep.* **3**, 1532 (2013).
- 751 29. Shaw, J. & Renzaglia, K. Phylogeny and diversification of bryophytes. *Am. J. Bot.*
752 **91**, 1557–1581 (2004).
- 753 30. Wu, Z. *et al.* Multiple transcriptional factors control stomata development in rice.
754 *New Phytol.* **223**, 220–232 (2019).

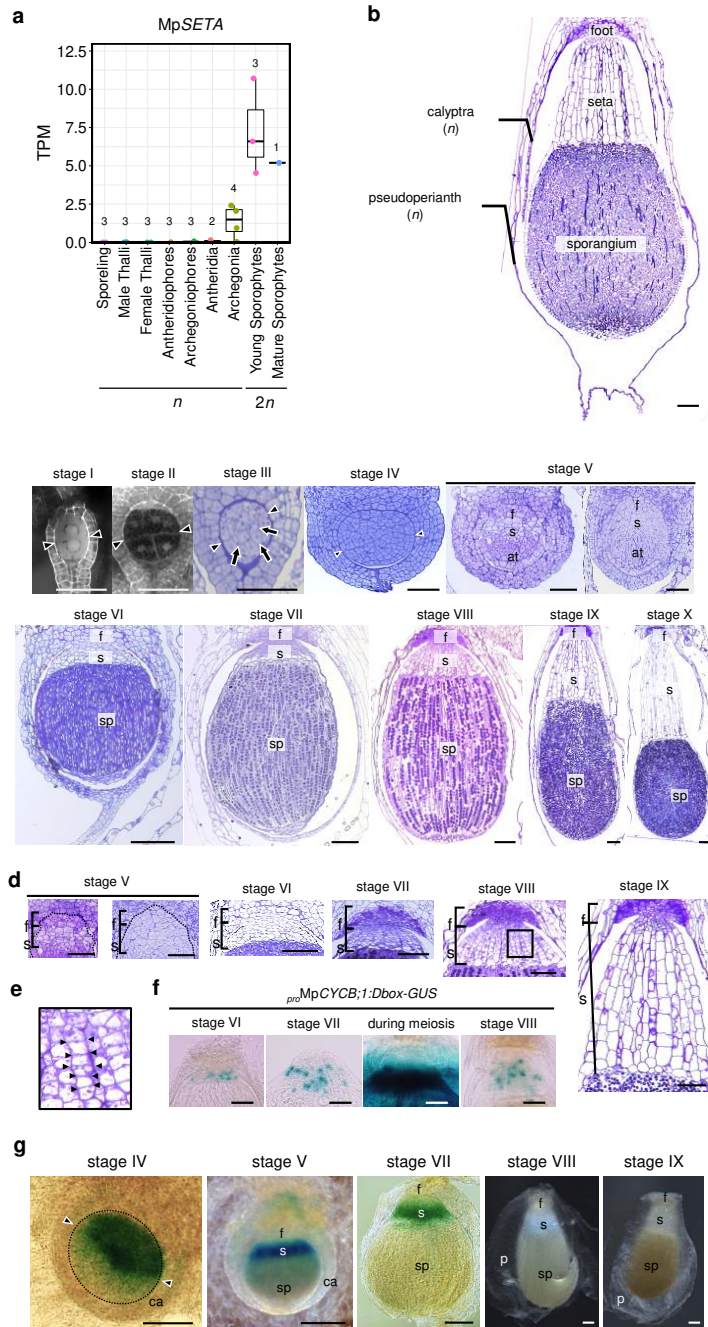
- 755 31. Sugano, S. S. *et al.* Efficient CRISPR/Cas9-based genome editing and its
756 application to conditional genetic analysis in *Marchantia polymorpha*. *PLoS One*
757 **13**, e0205117 (2018).
- 758 32. Sugano, S. S. *et al.* CRISPR/Cas9-mediated targeted mutagenesis in the liverwort
759 *Marchantia polymorpha* L. *Plant Cell Physiol.* **55**, 475–481 (2014).
- 760 33. Chinnusamy, V. *et al.* ICE1: A regulator of cold-induced transcriptome and
761 freezing tolerance in *Arabidopsis*. *Genes Dev.* **17**, 1043–1054 (2003).
- 762 34. Davies, K. A. & Bergmann, D. C. Functional specialization of stomatal bHLHs
763 through modification of DNA-binding and phosphoregulation potential. *Proc. Natl.*
764 *Acad. Sci. U. S. A.* **111**, 15585–15590 (2014).
- 765 35. MacAlister, C. A. & Bergmann, D. C. Sequence and function of basic helix-loop-
766 helix proteins required for stomatal development in *Arabidopsis* are deeply
767 conserved in land plants. *Evol. Dev.* **13**, 182–192 (2011).
- 768 36. Merced, A. & Renzaglia, K. S. Structure, function and evolution of stomata from
769 a bryological perspective. *Bryophyt. Divers. Evol.* **39**, 7–20 (2017).
- 770 37. Wei, D. *et al.* INDUCER OF CBF EXPRESSION 1 is a male fertility regulator
771 impacting anther dehydration in *Arabidopsis*. *PLoS Genet.* **14**, e1007695 (2018).
- 772 38. Higo, A. *et al.* Transcription factor DUO1 generated by neo-functionalization is
773 associated with evolution of sperm differentiation in plants. *Nat. Commun.* **9**, 5283
774 (2018).
- 775 39. Menand, B. *et al.* An ancient mechanism controls the development of cells with a
776 rooting function in land plants. *Science* **316**, 1477–1480 (2007).
- 777 40. Koshimizu, S. *et al.* *Physcomitrella* MADS-box genes regulate water supply and
778 sperm movement for fertilization. *Nat. Plants* **4**, 36–45 (2018).
- 779 41. Shirakawa, M. *et al.* FAMA is an essential component for the differentiation of
780 two distinct cell types, myrosin cells and guard cells, in *Arabidopsis*. *Plant Cell* **26**,
781 4039–4052 (2014).
- 782 42. Yasui, Y. *et al.* GEMMA CUP-ASSOCIATED MYB1, an ortholog of axillary
783 meristem regulators, is essential in vegetative reproduction in *Marchantia*
784 *polymorpha*. *Curr. Biol.* **29**, 3987–3995.e5 (2019).
- 785 43. Proust, H. *et al.* RSL class I genes controlled the development of epidermal
786 structures in the common ancestor of land plants. *Curr. Biol.* **26**, 93–99 (2016).
- 787 44. Puttick, M. N. *et al.* The interrelationships of land plants and the nature of the
788 ancestral embryophyte. *Curr. Biol.* **28**, 733–745.e2 (2018).
- 789 45. Renzaglia, K. S. & Garbary, D. J. Motile gametes of land plants: Diversity,
790 development, and evolution. *Crit. Rev. Plant Sci.* **20**, 107–213 (2001).
- 791 46. Sakakibara, K., Nishiyama, T., Deguchi, H. & Hasebe, M. Class 1 KNOX genes
792 are not involved in shoot development in the moss *Physcomitrella patens* but do
793 function in sporophyte development. *Evol. Dev.* **10**, 555–566 (2008).
- 794 47. Coudert, Y., Novák, O. & Harrison, C. J. A KNOX-cytokinin regulatory module
795 predates the origin of indeterminate vascular plants. *Curr. Biol.* **29**, 2743–2750.e5
796 (2019).
- 797 48. Li, M. & Sack, F. D. Myrosin idioblast cell fate and development are regulated by
798 the *Arabidopsis* transcription factor FAMA, the auxin pathway, and vesicular
799 trafficking. *Plant Cell* **26**, 4053–4066 (2014).
- 800 49. Shirakawa, M., Tanida, M. & Ito, T. The cell differentiation of idioblast myrosin
801 cells: Similarities with vascular and guard cells. *Front. Plant Sci.* **12**, 829541

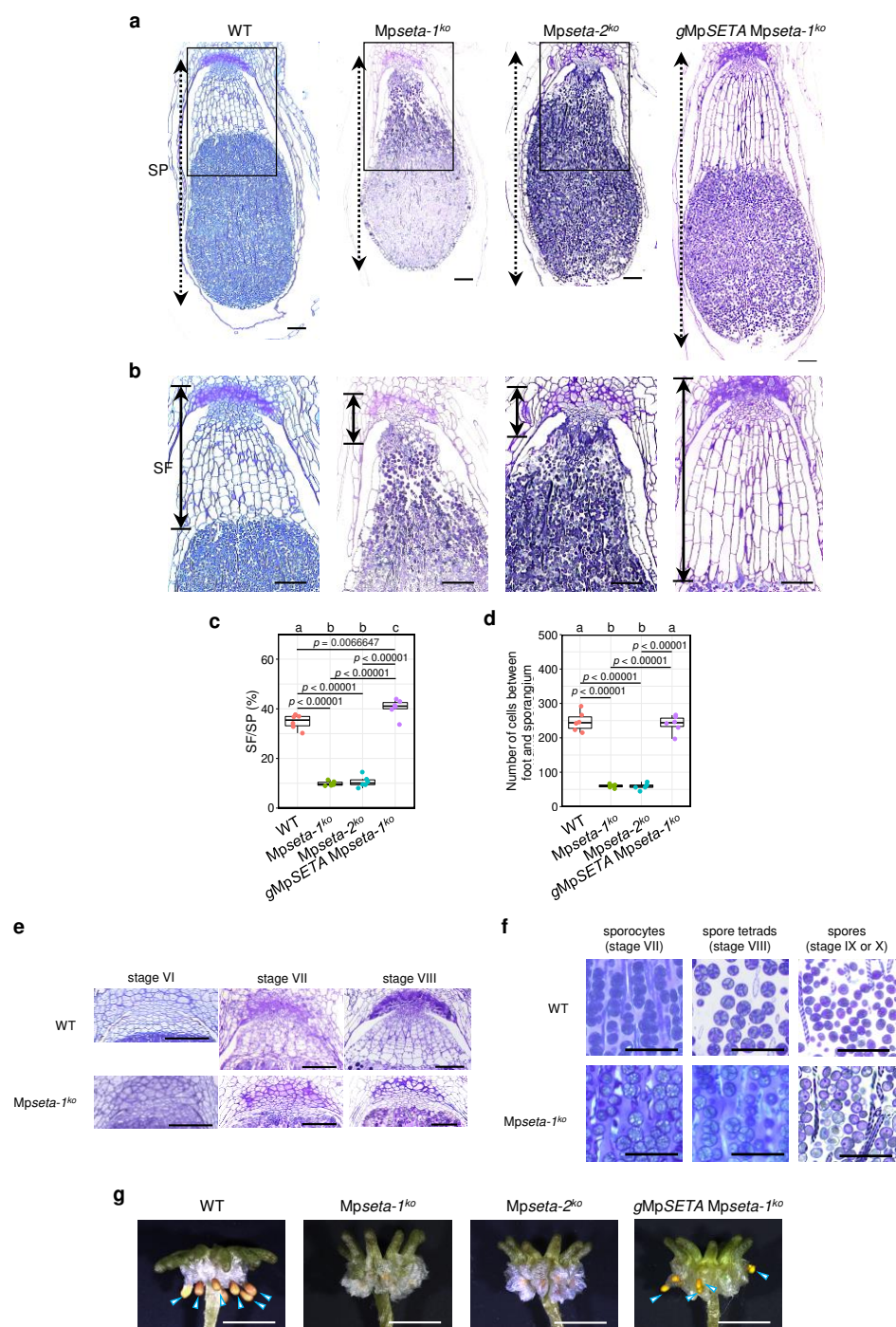
(2021).

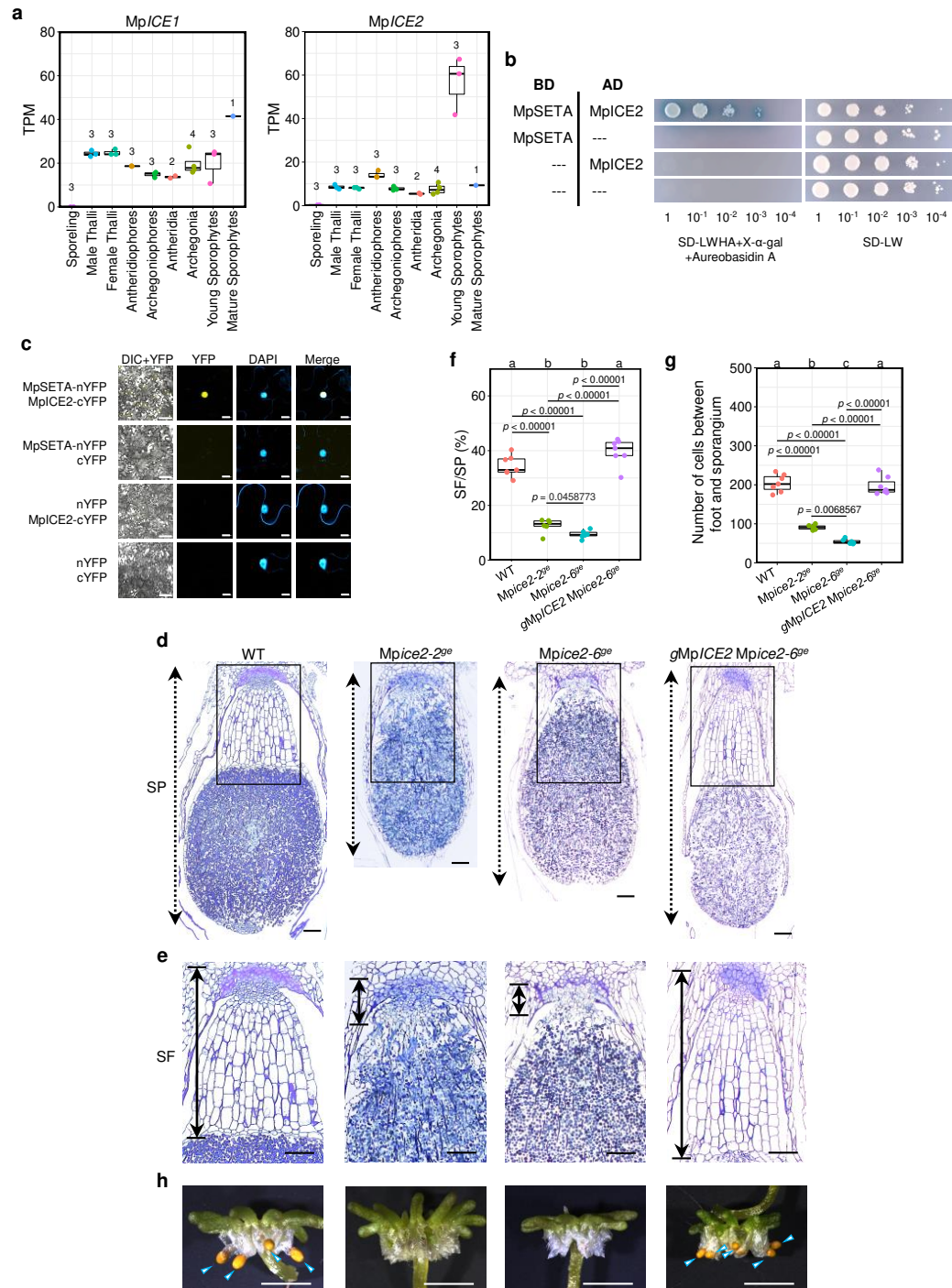
50. Li, F. W. *et al.* *Anthoceros* genomes illuminate the origin of land plants and the unique biology of hornworts. *Nat. Plants* **6**, 259–272 (2020).
51. Han, S. *et al.* MUTE directly orchestrates cell-state switch and the single symmetric division to create stomata. *Dev. Cell* **45**, 303–315.e5 (2018).
52. Lau, O. S. *et al.* Direct roles of SPEECHLESS in the specification of stomatal self-renewing cells. *Science* **345**, 1605–1609 (2014).
53. One Thousand Plant Transcriptomes Initiative. One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* **574**, 679–685 (2019).
54. Katoh, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* **9**, 286–298 (2008).
55. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
56. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
57. Nakagawa, T. *et al.* Improved gateway binary vectors: High-performance vectors for creation of fusion constructs in transgenic analysis of plants. *Biosci. Biotechnol. Biochem.* **71**, 2095–2100 (2007).
58. Nakamura, S. *et al.* Gateway binary vectors with the bialaphos resistance gene, bar, as a selection marker for plant transformation. *Biosci. Biotechnol. Biochem.* **74**, 1315–1319 (2010).
59. Clough, S. J. & Bent, A. F. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743 (1998).
60. Shimada, T. L., Shimada, T. & Hara-Nishimura, I. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J.* **61**, 519–528 (2010).
61. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
62. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
63. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
64. Li, B. & Dewey, C. N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
65. Ishizaki, K. *et al.* Development of gateway binary vector series with four different selection markers for the liverwort *Marchantia polymorpha*. *PLoS One* **10**, e0138876 (2015).
66. Matsushita, T., Mochizuki, N. & Nagatani, A. Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. *Nature* **424**, 571–574 (2003).
67. Tsuboyama, S., Nonaka, S., Ezura, H. & Kodama, Y. Improved G-AgarTrap : A highly efficient transformation method for intact gemmalings of the liverwort *Marchantia polymorpha*. *Sci. Rep.* **8**, 10800 (2018).
68. Ishizaki, K., Chiyoda, S., Yamato, K. T. & Kohchi, T. *Agrobacterium*-mediated transformation of the haploid liverwort *Marchantia polymorpha* L., an emerging model for plant biology. *Plant Cell Physiol.* **49**, 1084–1091 (2008).
69. Rövekamp, M., Bowman, J. L. & Grossniklaus, U. *Marchantia* MpRKD regulates

- 849 the gametophyte-sporophyte transition by keeping egg cells quiescent in the
850 absence of fertilization. *Curr. Biol.* **26**, 1782–1789 (2016).
- 851 70. Rossignol, P., Collier, S., Bush, M., Shaw, P. & Doonan, J. H. *Arabidopsis* POT1A
852 interacts with TERT-V(18), an N-terminal splicing variant of telomerase. *J. Cell*
853 *Sci.* **120**, 3678–3687 (2007).
- 854 71. Hino, T. *et al.* Two Sec13p homologs, AtSec13A and AtSec13B, redundantly
855 contribute to the formation of COPII transport vesicles in *Arabidopsis thaliana*.
856 *Biosci. Biotechnol. Biochem.* **75**, 1848–1852 (2011).
857

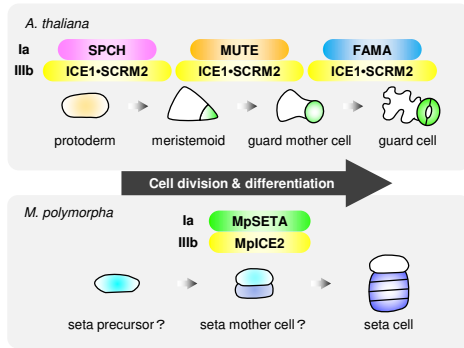




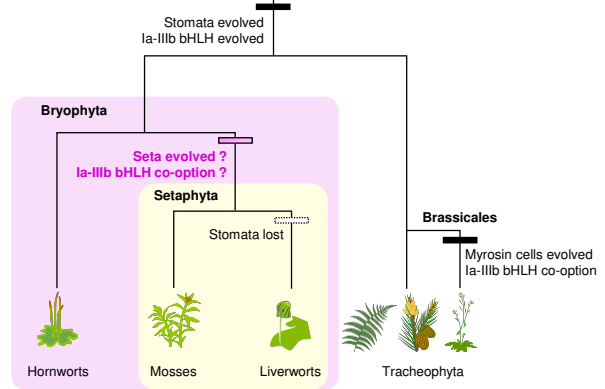




a



b

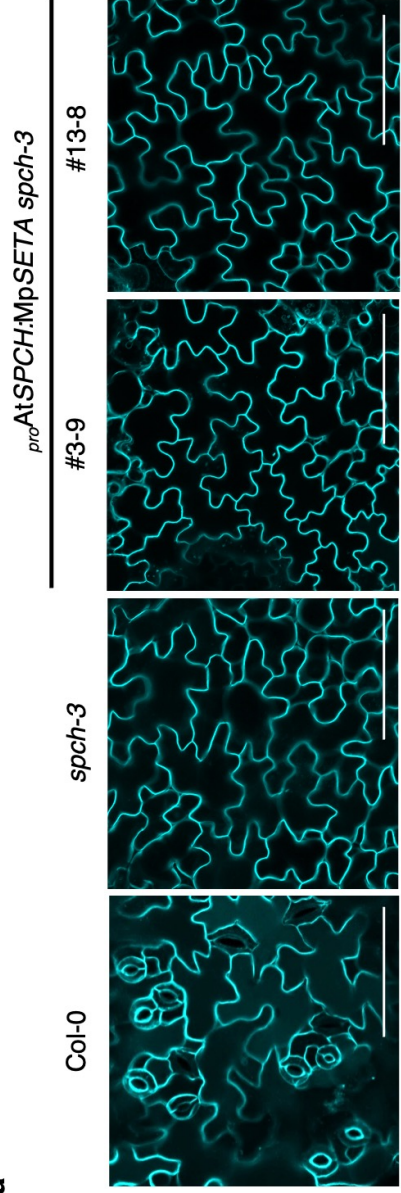


Protein	bHLH	SMF	PEST	Grey	Total
MpSETA	268	327	385	461	1441
PpSMF1	426	485	591	669	2171
PpSMF2	318	377	483	561	1739
AtSPCH	101	160	286	363	910
AtMUTE	59	115	192	1	367
AtFAMA	196	255	316	393	1160

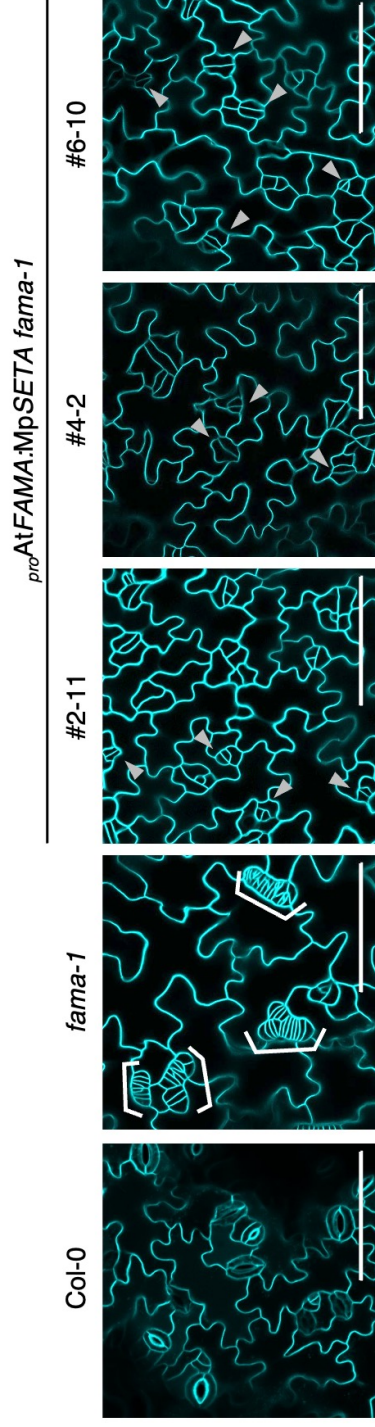
[illegible][illegible]

MpSETA	Q	A	L	Y	K	L	Q	A	E	I	-	S	M	R	V	N	-	-	-	-	-	-	-	A	K	V	I	A	S	R	I	H	E	S	
LcSETA1	K	T	G	S	Y	K	F	Q	A	K	V	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	R	F	V	L	A	L	T	R	I
LcSETA2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
PpSMF1	T	V	L	Y	S	F	H	V	Q	I	-	S	L	H	C	R	-	L	-	N	V	D	E	V	A	A	A	L	H	O	T	-	-	-	-
PpSMF2	T	V	L	F	S	F	H	V	H	M	-	G	L	L	C	R	-	M	-	S	V	K	E	I	A	T	V	L	H	G	T	-	-	-	-
AtSPCH	T	M	L	N	S	F	T	I	K	I	-	G	I	E	C	Q	-	L	-	S	A	E	E	L	A	Q	Q	I	Q	T	-	-	-	-	
AtMUTE	T	V	L	Y	F	F	V	V	K	I	-	G	L	E	C	H	-	L	-	S	L	E	E	L	T	L	E	V	Q	K	S	-	-	-	
AtFAMA	T	V	L	Y	S	F	N	V	K	I	-	T	S	E	T	R	-	F	-	T	A	E	D	I	A	S	S	I	Q	I	-	-	-	-	
AtbHLH57	F	V	I	Y	S	F	N	L	K	M	-	E	D	G	C	K	-	L	G	-	S	A	D	E	I	A	T	A	V	H	Q	-	-	-	
AtbHLH67	S	V	S	Y	S	F	N	L	K	M	-	E	D	E	C	D	-	L	E	-	S	A	D	E	I	T	A	A	V	H	R	I	-	-	
AtbHLH70	S	V	S	Y	S	F	N	L	K	M	-	E	D	E	C	N	-	L	G	-	S	A	D	E	I	T	A	A	I	R	Q	I	-	-	
AtbHLH71	Y	A	I	Y	S	I	A	K	V	-	E	E	S	C	Q	-	L	S	-	S	V	D	D	I	A	G	A	V	H	M	T	-	-	-	
AtbHLH94	S	I	L	Y	S	I	S	V	R	V	-	E	E	G	S	Q	-	L	N	-	T	V	D	D	I	A	T	A	L	N	Q	T	-	-	
AtbHLH96	S	V	L	Y	S	I	S	V	K	V	-	E	E	G	S	Q	-	L	N	-	T	V	E	D	I	A	A	V	N	Q	I	-	-		
AtbHLH99	M	S	L	F	T	F	S	V	K	V	-	E	A	D	C	Q	-	L	T	P	S	G	N	E	V	A	N	T	V	H	E	V	-	-	
AtZOU	W	S	T	Y	V	I	Q	A	Q	V	-	N	P	S	C	E	N	Q	L	L	G	N	G	L	G	V	V	D	V	F	K	Q	V	-	-
MpBHLH38	T	I	M	Y	M	V	H	A	Q	T	-	E	D	Q	L	G	-	-	-	-	N	S	D	M	L	N	I	A	L	Q	D	L	-	-	
MpBHLH24	T	V	F	H	C	L	H	V	M	T	-	P	E	P	G	C	D	-	F	-	P	K	E	A	L	Q	T	A	L	Q	N	-	-	-	

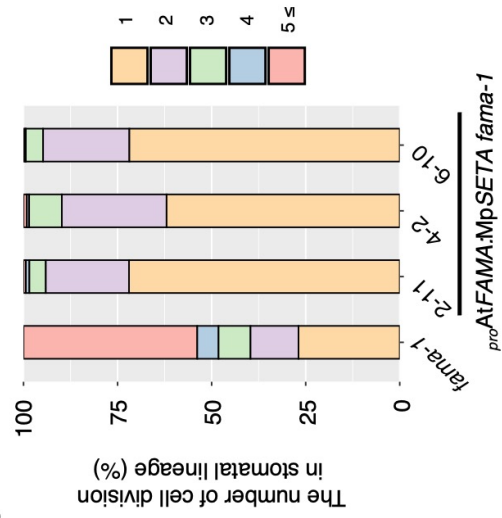
a



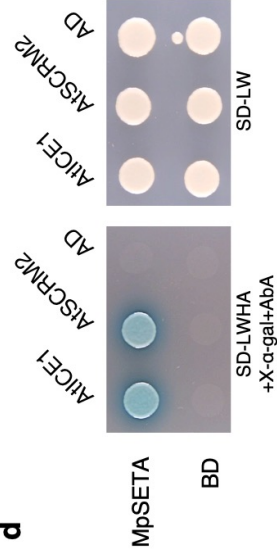
b



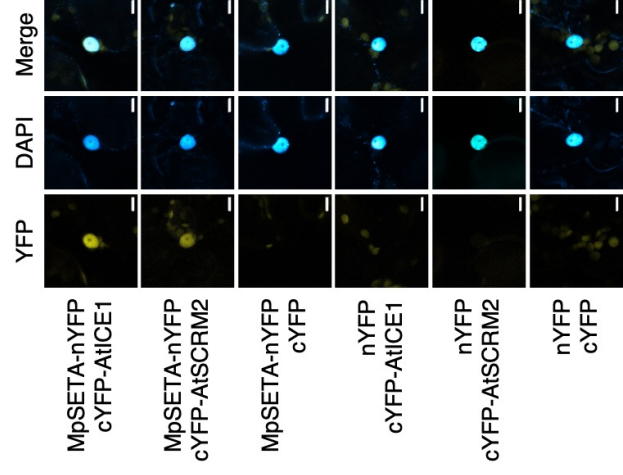
c



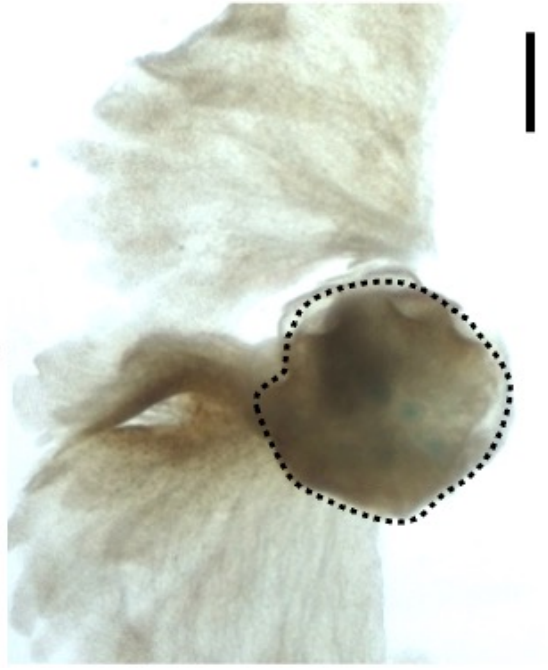
d



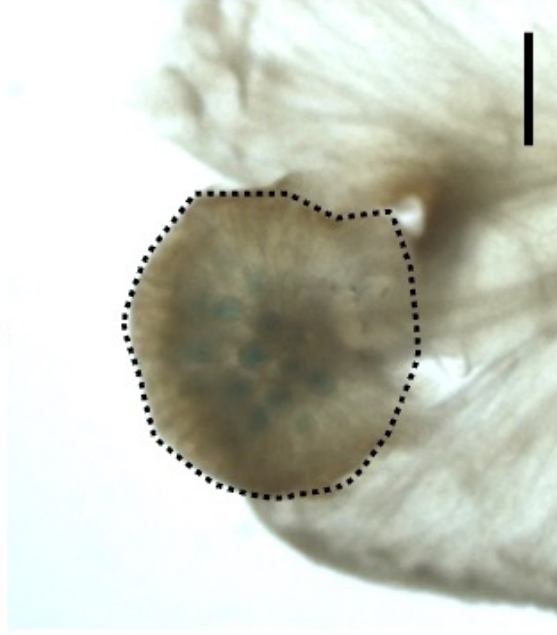
e



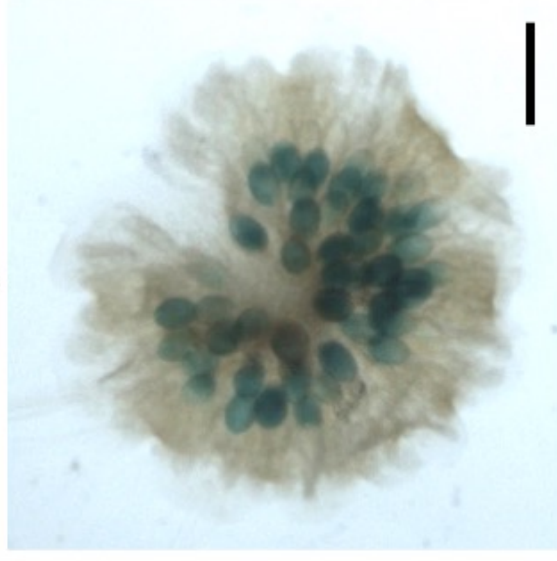
stage 1



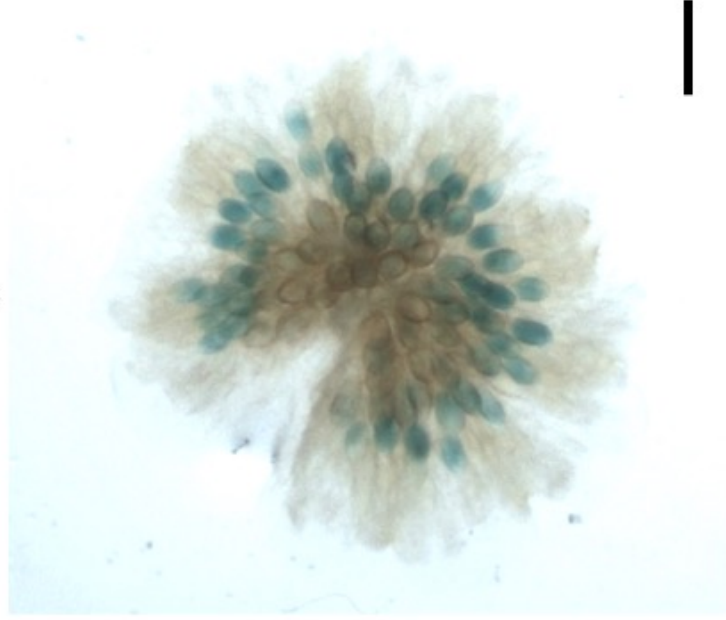
stage 2



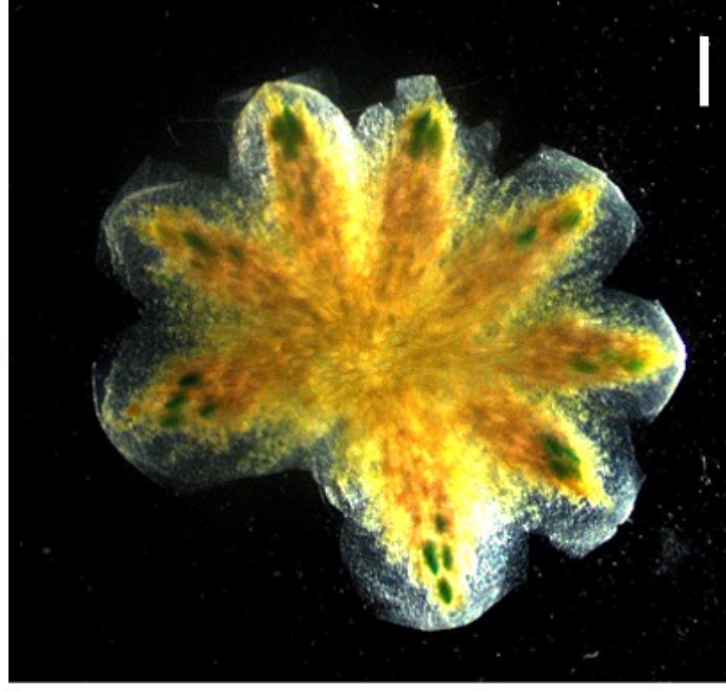
stage 3

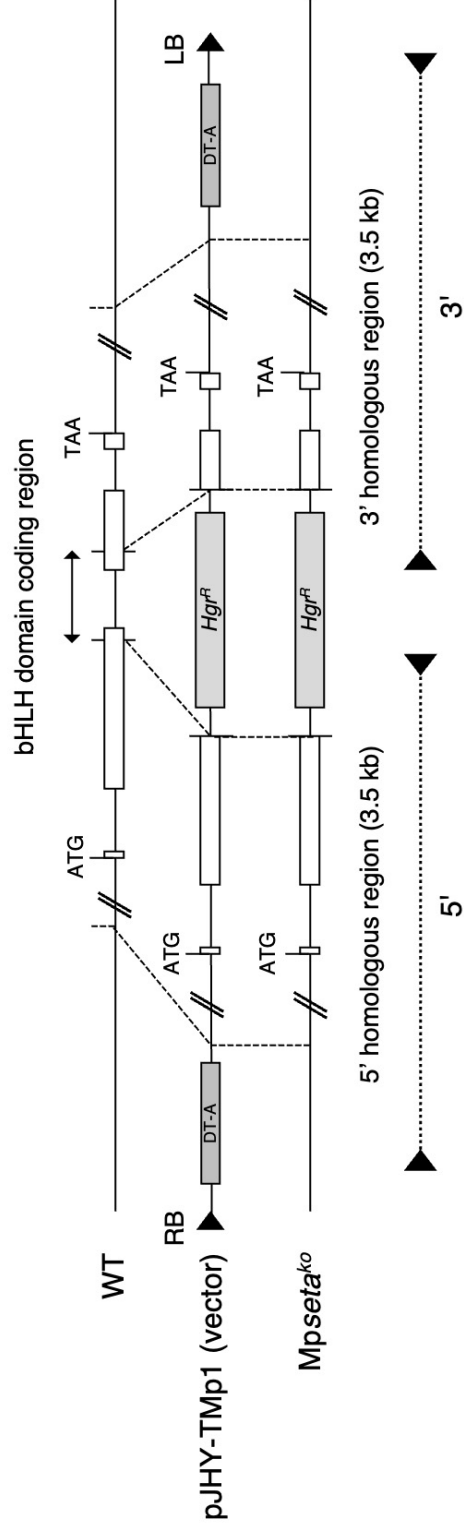
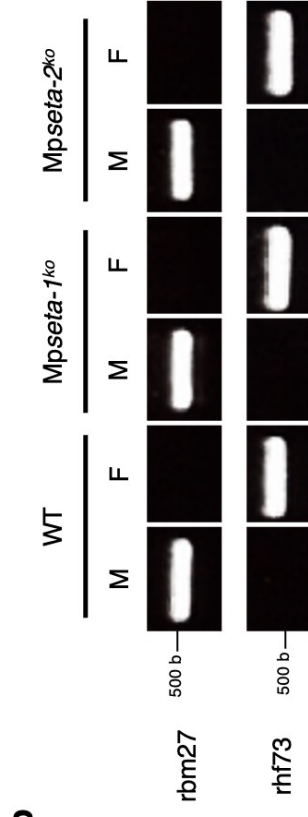
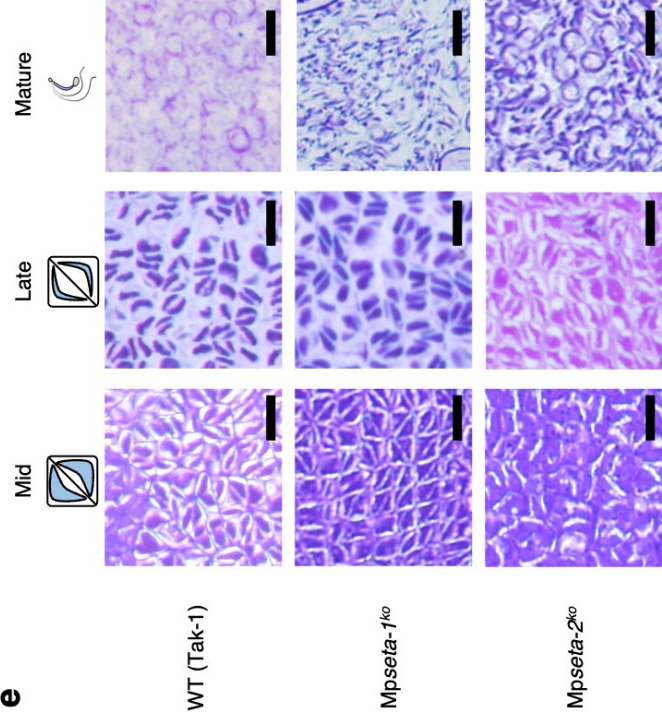
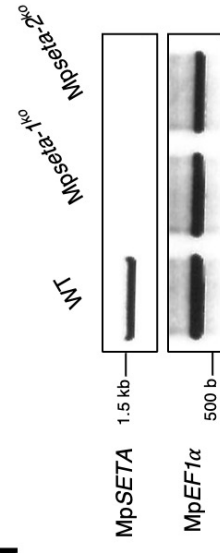


stage 4



stage 5



a**b****e****d**

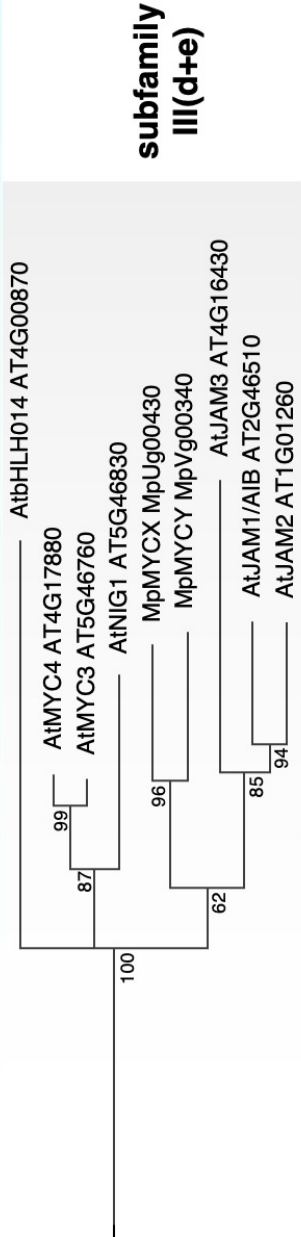
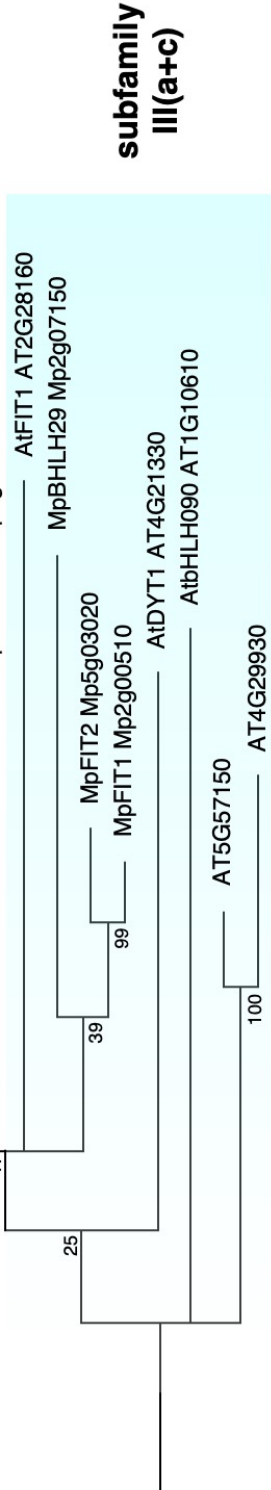
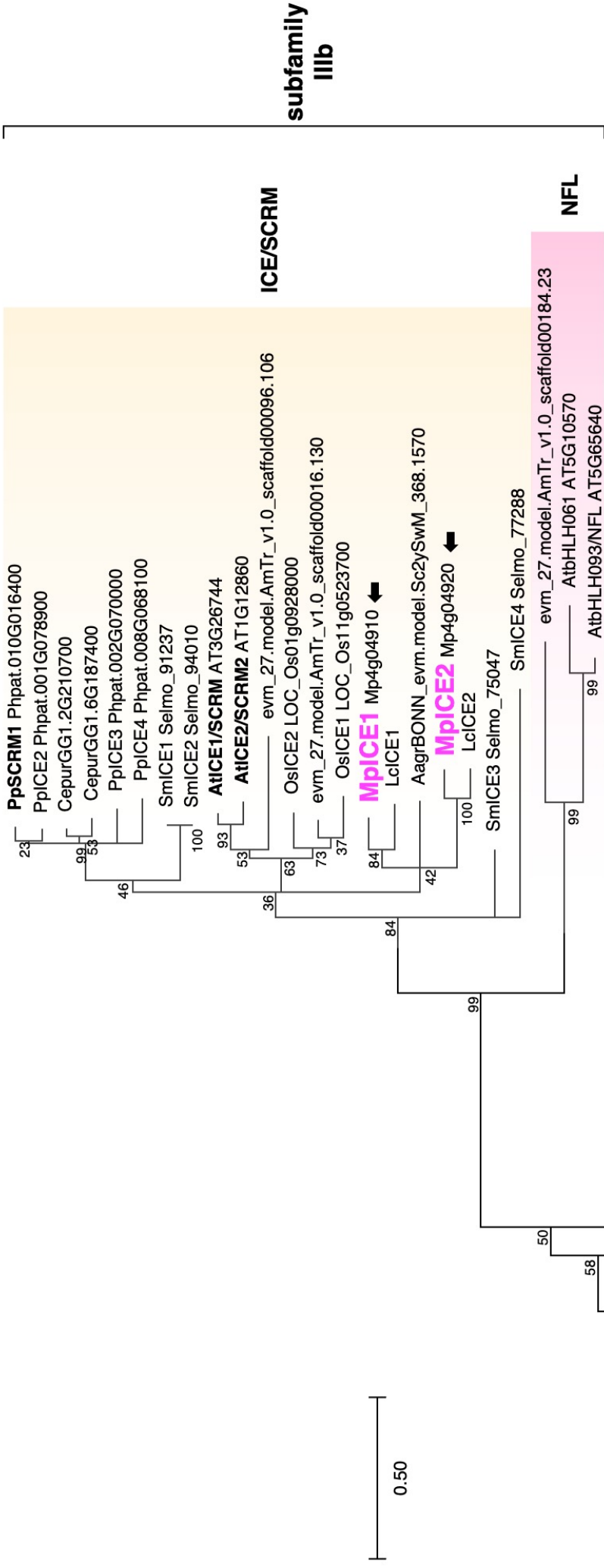
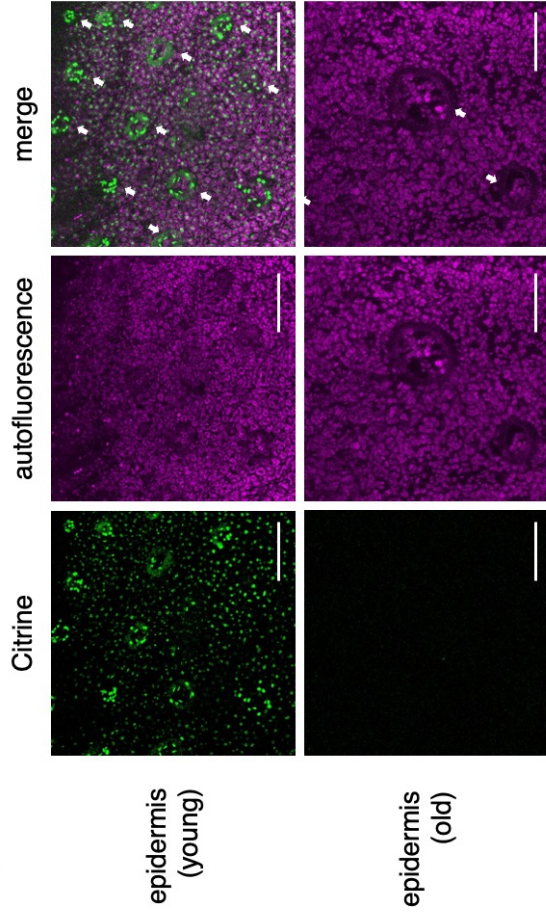
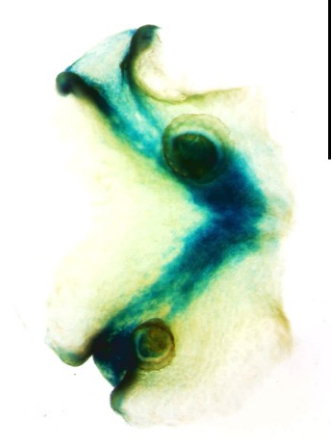
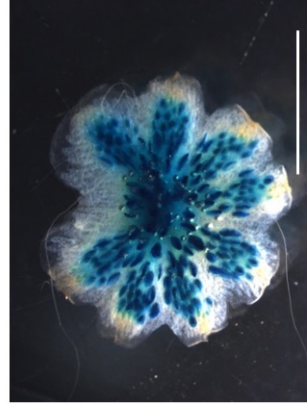
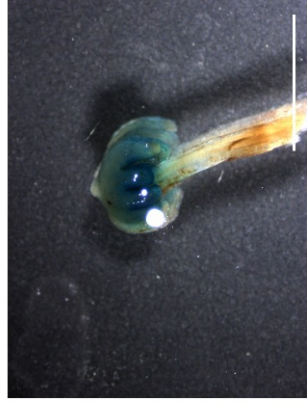
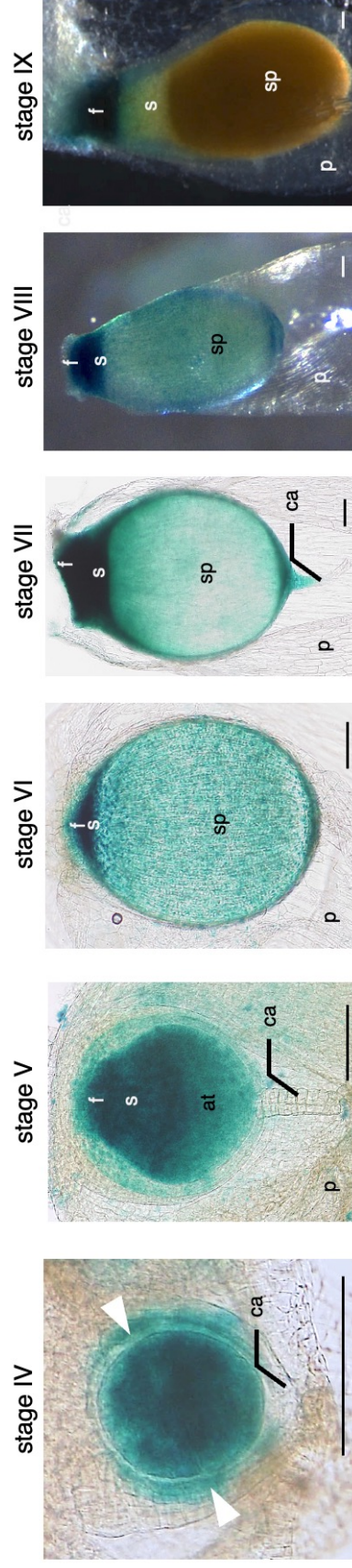


Figure 1: Phylogenetic relationships and domain architecture of ICE1 and SCRM proteins. The figure shows six protein sequences: MpICE1, MpICE2, PpSCRM1, AtICE1/SCRM, and AtICE2/SCRM2. Each sequence is represented by a horizontal bar with a scale from 1 to the total length. Domains are color-coded: blue for bHLH and yellow for ACT-like. The domain architecture is as follows: MpICE1 (bHLH: 504-560, ACT-like: 617-695), MpICE2 (bHLH: 499-555, ACT-like: 587-660), PpSCRM1 (bHLH: 287-343, ACT-like: 389-445), AtICE1/SCRM (bHLH: 305-361, ACT-like: 410-467), and AtICE2/SCRM2 (bHLH: 265-321, ACT-like: 347-423). A legend at the bottom indicates that blue represents bHLH and yellow represents ACT-like domains.

[illegible]

MpICE1	LNIHMFCSRRPGLLLSTMRALDGLGLEVQQAVVSCFNGFALDVFRAEQLQDNE-IGSEEVRSVLL
MpICE2	LNIHMLCARKPGLLLSIMRALDGLGLEVEQGVSFCFNGFALDI FRAEQLSGLD-VGPEEIRQALL
LcICE1	INIHMFSTKRPGLLLSTMRALEGLGI EVQQAVVSCFNGFALDI FRAE-----
LcICE2	LNIHMLCARKPGLLISVMRALDGLGFVEVQGVVSCFNGFALDI FRAEQLKGMD-VTPPEER EALL
PpSCRM1	LNIHMFCARRPGLLLSTVKALDALGLDVQQAVISCFCNGFALDLFRAEAKDVD--VGPEEIKAVLL
PpICE2	LNIHMFCARRPGLLLSTVRALDALGLDVQQAVISCFCNGFALDLFRAEAKDVD--VGPEEIKAVLL
PpICE3	LNIHMFCARRPGLLLSTVRALDALGLDVQQAVISCFCNGFALDLFRAEAKDAD--VEPDEIKAVLL
PpICE4	LNIHMFCARRRGLLLSTVRALDALDL DVQQAVISCFCNGFALDLFRAEVSRAR--PTPKP-----
AtICE1/SCRM	VNIHMF CGRRPGLLLATMKALDNLGLDVQQAVISCFCNGFALDVFRAEVIFYSQ-LKG-----
AtICE2/SCRM2	VNIHMF C GRRPGLLLSTMRALDNLGLDVQQAVISCFCNGFALDVFRAEV I-----
AthHLH061	THIDICCP TKPGLVVSTVSTLET LGLEIEQCVI SCFSDFS LQA S CF EVGEQRYMVT SEATK QALI
AthHLH093/NFL	TRVDICSP KPGLLLSTVN TLET LGLEIEQCVI SCFSDFS LQA S CS EGAEQRDFIT SEDIK QALF
AtAMS	FFVKVI CEY K PGGFTRLMEALDSLGL E VTNANTTRYLSLVSNVFKVEKNDNEM-VQA EHVRNSLL
MpBHLH11	FTVRII SEKSEG MFLGVLRIMDEYGFDIL HANIITHRGIALHVFR AEIRDKKT-PTIEQVKESLL

b**a****c****d****e**

a

Mp/CE2 (Mp4g04920)



GCAAAAGGCTTGCCTGCGAAGATC TGGCAGAA

GCAAAAGGCTTGCCTGCGAAGATC NNN TGGCAGAA

WT (Tak-1)

WT (Tak-2)

Mpice2-2ge (Male)

Mpice2-2ge (Female)

GCAAAAGGCTTGCCTGCGAAGATC TCA TGGCAGAA
GCAAAAGGCTTGCCTGCGAAGATC TCA TGGCAGAA
GCAAAAGGCTTGCCTGCGAAGATC TCA TGGCAGAA
GCAAAAGGCTTGCCTGCGAAGATC TCA TGGCAGAA
GCAAAAGGCTTGCCTGCGAAGATC TCA TGGCAGAA

ACAGAGCATCAAT CCT TGGGGA TCCCCATCAATCTGGGACGCAATCGA

ACAGAGCATCAAT NNN TGGGGA NNNNNNNNNNNN CGCAATCGA

WT (Tak-1)

WT (Tak-2)

Mpice2-6ge (Male)

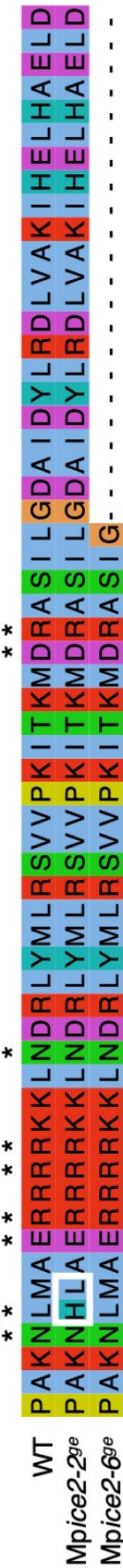
Mpice2-6ge (Female)

ACAGAGCATCAAT CCT TGGGGA -- -- -- -- -- CGCAATCGA
ACAGAGCATCAAT CCT TGGGGA -- -- -- -- -- CGCAATCGA
ACAGAGCATCAAT -- -- TGGGGA TCCCCATCAATCTGGGACGCAATCGA
ACAGAGCATCAAT -- -- TGGGGA TCCCCATCAATCTGGGACGCAATCGA

3 bp

17 bp

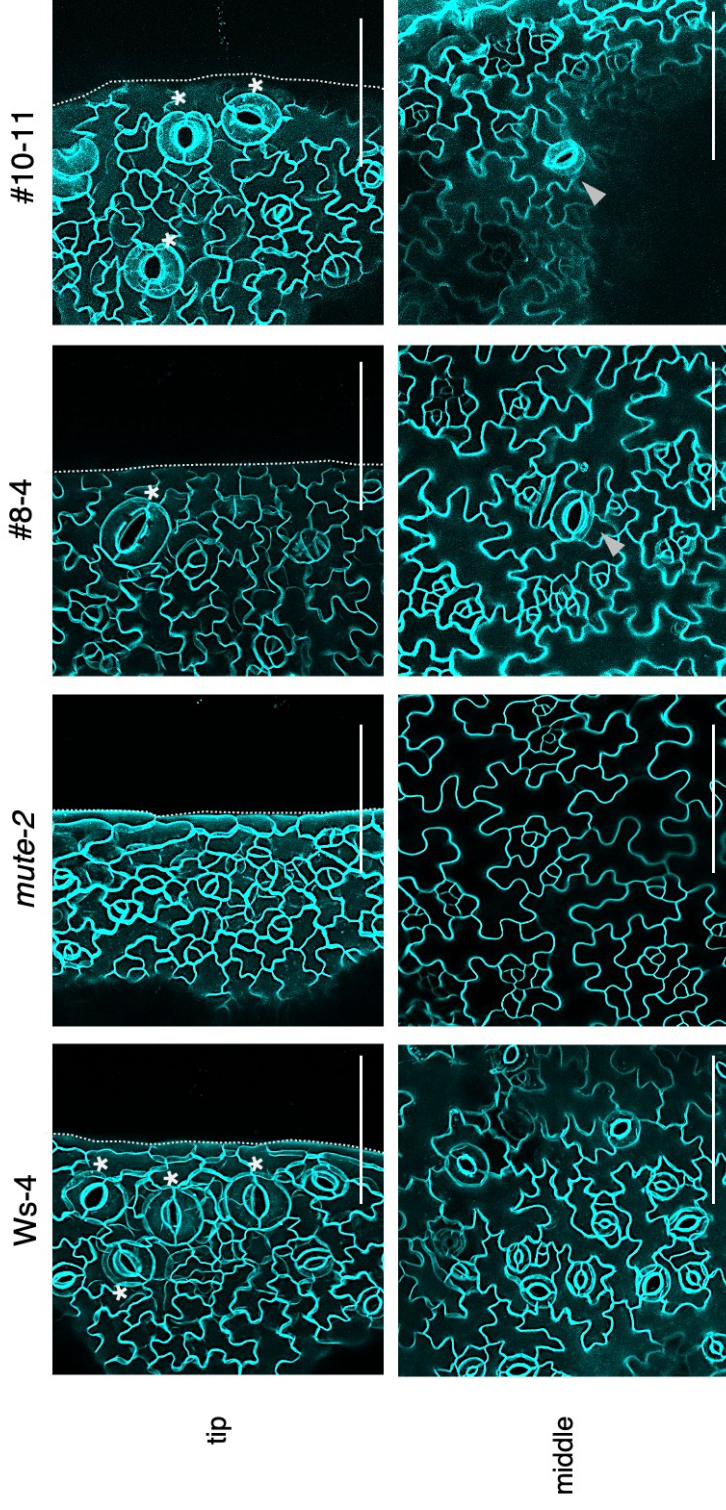
b



a

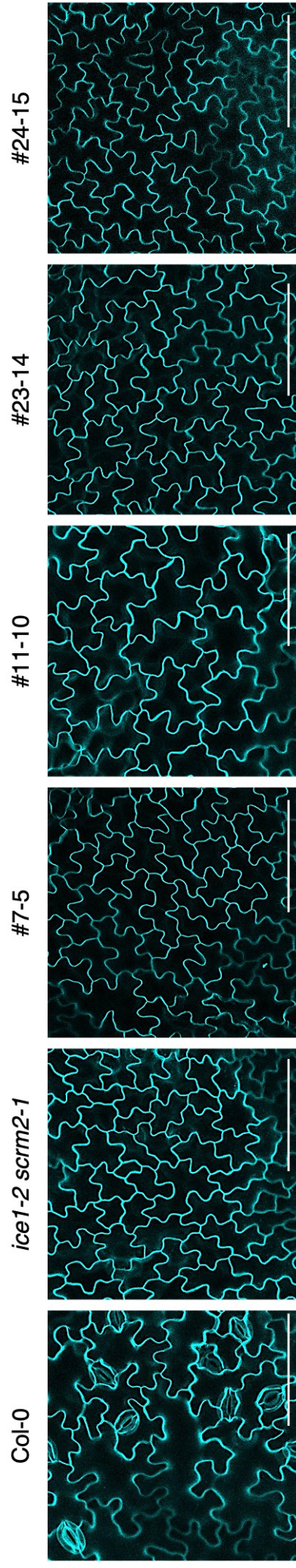
pro^{35S:MpICE2}

pro^{AtMUTE:MpSETA} *mute-2*



b

pro^{AtICE1:MpICE1} *ice1-2 scrm2-1*



Supplementary Table 1 | Cloning primers used in this study.

Plasmid	Primer Name	Sequence (5'>3')	Source
pENTR1A-MpSETA	MpSETA_CDS_ENTR_F	atccggtaccgaattcgcATGGACGGTGTGTTCTGGACGATG	this study
	CDS	MpSETA_CDS_ENTR_R	gtgcggccgcgaattctaACCCATGTTCTCAAATTGAAGATTG
pENTR1A-MpICE2	MpICE2_CDS_ENTR_F	atccggtaccgaattcgcATGGCTTCGACCAGAGT	this study
	CDS	MpICE2_CDS_ENTR_R	gtgcggccgcgaattctaCTGAAGAATCACATTTTGCTCT
pENTR1A- <i>pro</i> MpSETA	MpSETApro_ENTR_F	aaccaattcagtcgacCAAGCCAGGTGAAAAAATC	this study
	MpSETApro_ENTR_R	gtgcggccgcgaattCCGAAGCGATTGGACCAGTAT	this study
pENTR1A- <i>pro</i> MpICE2:Citrine- GUS-NLS	MpICE2pro_ENTR_F	atccggtaccgaattcgcACATTCTACTTGCAGCACTGA	this study
	MpICE2pro_Cit_R	gcccttgctcaccatGAGAGAGTGAGCGCCTTCG	this study
	Citrine_proMpICE2_F	GGCGCTCACTCTCTCatggtgagcaagggcgag	this study
	Citrine_GUS_R	acgtaacatctgcagctgttacgctcgtccatgc	this study
	GUS_Citrine_F	gacgagctgtacaagctgcagatgttacgtcctgt	this study
	GUS-NLS_ENTR_R	gtgcggccgcgaattctatcctccaacctttcttct	this study
pENTR1A-gMpSETA	MpSETApro_ENTR_F & MpSETA_CDS_ENTR_R	see above	this study
pENTR1A-gMpICE2	MpICE2pro_ENTR_F & MpICE2_CDS_ENTR_R	see above	this study
pDEST-GBKT-Amp ^r	AmprPro_Sfol	TGTCAGCGCAGGGGCGCCCGCGGAACCCCTATTTGTTT	this study
	AmprCDS_Sfol	CAAAAAGAACCGGGCGCCTTACCAATGCTTAATCAGTGAGGC	this study
pENTR1A-AtICE1	AtICE1_CDS_ENTR_F	aaccaattcagtcgacATGGGTCTTGACGGAAACAA	this study
	AtICE1_CDS_ENTR_R	aagctgggtctagatatccTCAGATCATACCAGCATACCCTG	this study
pENTR1A-AtSCRM2	AtSCRM2_CDS_ENTR_F	aaccaattcagtcgacATGAACAGCGACGGTGTGTTG	this study
	AtSCRM2_CDS_ENTR_R	aagctgggtctagatatccTCAAACCAAACCAGCGTAACC	this study
pENTR 5'/TOPO- <i>pro</i> AtMUTE	AtMUTEpro_F	CAGCAATTTGAAAAATCCA	this study
	AtMUTEpro_R	gatacttaattgatcaagata	this study
pENTR 5'/TOPO- <i>pro</i> AtFAMA	AtFAMApro_F	cttccaaacttcagtgtgata	this study
	AtFAMApro_R	tgctattcgtggtagttgata	this study
pENTR1A- <i>pro</i> AtSPCH:MpSETA	SPCHpro_ENTR_F	aaccaattcagtcgacAGATCATCACTGCGATAAGG	this study
	SPCHpro_Mp_R	AACAACACCGTCCATCGTGATTAGAGATATATCCT	this study
	MpSETA_SPCHpro_F	ATCTCTAATCACGATGGACGGTGTGTTCTGGACGATG	this study
pENTR1A- <i>pro</i> AtICE1:MpICE1	proICE1_ENTR_F_Sal1	aaccaattcagtcgacGGACCACCGTCAATAACATCG	this study
	proICE1_MpBHLH16_R	TCTGGTTGTAGAAGTCATCATGCCAAAGTTGACACCTTTACC	this study
	MpBHLH16_proICE1_F	gtaaagggtgcaactttggcATGATGACTTCTACAACCAGA	this study
	MpBHLH16_CDS_ENTR_R	gtgcggccgcgaattctaCTGAAGGGCACTGACGTTA	this study
pENTR1A- <i>pro</i> AtICE1:MpICE2	proICE1_MpBHLH17_R	CTGACTCTGGTCTGAAGCCATGCCAAAGTTGACACCTTTACC	this study
	MpBHLH17_proICE1_F	gtaaagggtgcaactttggcATGGCTTCGACCAGAGT	this study
pJHY-Tmp1- MpSETAko	MpSETA-KO-Pac-F	ctaaggtagcgattaatACTCCGAATTTAAACGAT	this study
	MpSETA-KO-Pac-R	gcccgggcaagcttaatCCGGTAGCTTTTATTTCGC	this study
	MpSETA-KO-Asc-F	taaactagtggcgcgCTGCAAGATCAAAAGAAT	this study
	MpSETA-KO-Asc-R	ttatccctaggcgcgTGACAATGCTTCGATCAC	this study

Supplementary Table 2 | Genotyping, RT-PCR, and genome-editing primers used in this study.

Purpose	Name	Sequence (5'>3')	Source
For genotyping sequencing (<i>Mpseta^{ko}</i>)	MpSETA_KO_F	AACAAAAAAGAAGCGAATAAAAGCTACCGG	this study
	MpSETA_KO_R	AGCGATGCTTGATTCTTTTGATCTTGCAG	this study
	MpSETA_5'genomic_F	TGTTTTCACTCTCTCTATGACTTTTCAAGT	this study
	MpSETA_3'genomic_R	TATCTGGGTTTTCTAGAGCTCATATAAACT	this study
	P1R	GAAGGCTTCTGATTGAAGTTTCCTTTTCTG	Ishizaki et al. ¹
	H1F	GTATAATGTATGCTATACGAAGTTATGTTT	Ishizaki et al. ¹
For sex diagnosis	rbm27-F	ACTTTTGCAACAGCGACTTC	Fujisawa et al. ²
	rbm27-R	GCCTGCAATATAGCCTTCAA	Fujisawa et al. ²
	rhf73-F	GAACCCGAAACTCAGGTTTT	Fujisawa et al. ²
	rhf73-R	ATAACAGCCAAACGGATCAA	Fujisawa et al. ²
For RT-PCR	MpSETA_RT_F	ATGGACGGTGTGTTCTGGACGATG	this study
	MpSETA_RT_R	ACCCATGTTCTCAAATTGAAGATTG	this study
	MpEF1 α _F	AGGTTGTCACCATGGGAAAGGAGA	Rövekamp et al. ³
	MpEF1 α _R	TCACACGCTTGTCATACCTCCCA	Rövekamp et al. ³
For genome editing (<i>Mpice2^{ge}</i>)	MpICE2_sgRNA1_F	CTCG CTTGCCTGCGAAGAATCTCA	this study
	MpICE2_sgRNA1_R	AAAC TGAGATTCTTCGCAGGCAAG	this study
	MpICE2_sgRNA2_F	CTCG GGACAGAGCATCAATCCTTG	this study
	MpICE2_sgRNA2_R	AAAC CAAGGATTGATGCTCTGTCC	this study

1. Ishizaki, K., Johzuka-Hisatomi, Y., Ishida, S., Iida, S. & Kohchi, T. Homologous recombination-mediated gene targeting in the liverwort *Marchantia polymorpha* L. *Sci. Rep.* **3**, 1–6 (2013).
2. Fujisawa, M. *et al.* Isolation of X and Y chromosome-specific DNA markers from a liverwort, *Marchantia polymorpha*, by representational difference analysis. *Genetics* **159**, 981–985 (2001).
3. Rövekamp, M., Bowman, J. L. & Grossniklaus, U. *Marchantia* MpRKD Regulates the Gametophyte-Sporophyte Transition by Keeping Egg Cells Quiescent in the Absence of Fertilization. *Curr. Biol.* **26**, 1782–1789 (2016).